An Improved ResNet Architecture for Accurate Patch-Level HER2 Overexpression Classification in Breast Cancer Tissue Images

Shubhangi Joshi
Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Lavale, Pune, Maharashtra, India
shubhangi.joshi.phd2020@sitpune.edu.in

Pallavi Chaudhari
Department of Pathology, Symbiosis Medical College for Women & Symbiosis University Hospital & Research Centre, Symbiosis International (Deemed University), Pune, Maharashtra, India
pallavi.chaudhari@smcw.siu.edu.in

Anupkumar Bongale
Department of Artificial Intelligence and Machine Learning, Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Lavale, Pune, Maharashtra, India
anupkumar.bongale@sitpune.edu.in (corresponding author)

Deepak Dharrao
Department of Computer Science and Engineering, Symbiosis Institute of Technology, Pune Campus, Symbiosis International (Deemed University), Lavale, Pune, Maharashtra, India
deepak.dharrao@sitpune.edu.in

Vivek Dugad
Department of Pathology, Symbiosis Medical College for Women & Symbiosis University Hospital & Research Centre, Symbiosis International (Deemed University), Pune, Maharashtra, India
vivek.dugad@smcw.siu.edu.in

Anand Bhosale
Department of Pathology, D.Y. Patil Deemed to be University, India | School of Medicine and Pushpalata, D.Y. Patil Hospital, Ambi, Pune, Maharashtra, India
anandbhosale2007@gmail.com

Received: 14 May 2025 | Revised: 13 June 2025, 1 July 2025, and 11 July 2025 | Accepted: 13 July 2025

 DOI: https://doi.org/10.48084/etasr.12140

ABSTRACT

With the advent of artificial intelligence, various computer-aided diagnostic systems are being developed to assist medical professionals. Deep learning techniques powered by convolutional neural networks seem promising for obtaining new insights into the onco-histopathology domain. Breast cancer is confirmed by histopathological analysis of Hematoxylin and Eosin (H&E)-stained breast tissue images. Finding the molecular subtype of breast cancer using Immunohistochemistry (IHC)-stained breast tissue is essential to decide on the correct treatment plan for a breast cancer patient. IHC staining is an expensive process that is very time-consuming and involves intra- and inter-observer subjectivity. This work attempts to find the Human Epidermal growth factor Receptor Two (HER2) molecular subtype from H&E-stained tissue images instead of using IHC-stained tissues. H&E-stained tissue image data from two separate sources are used to predict HER2 status. This study aimed to improve the accuracy of HER2 overexpression classification by modifying the architecture of the ResNet50 model by cascading it with a squeeze and

excitation block and a depth-wise separable convolutional layer. The dataset comprises a combination of tissue image patches from a publicly available Warwick dataset and a real-world dataset collected from a hospital in Pune, India. The dataset is preprocessed and split into 60% train, 20% validation, and 20% test subsets. The proposed architecture with a modified ResNet50 network achieves the best patch-level HER2 classification accuracy of 98.04%, with class-specific test accuracy results for HER2 negative, HER2 equivocal, and HER2 positive scores being 97.73%, 99.70%, and 98.93%, respectively.

Keywords-breast cancer; HER2; histopathology; H&E-staining; ResNet50; SE blocks; patches

I. INTRODUCTION

Breast cancer occurs due to rapid and abnormal cell growth in breast tissue and must be detected early. A delay in cancer diagnosis can cause cancer to spread to other organs, which can result in fatality. To confirm the presence of breast cancer, morphological analysis of breast tissue using Hematoxylin and Eosin (H&E) staining is the gold standard [1]. Histopathologists analyze the morphology of H&E-stained breast tissue to confirm the presence of breast cancer, determine its invasiveness, and classify the type of malignancy. Breast cancer has several molecular subtypes, such as luminal A, luminal B, HER2-positive, and triple-negative breast cancer. These molecular subtypes suggest different treatment plans and indicate different prognostic information. These molecular subtypes depend on hormonal changes in Estrogen (ER), Progesterone (PR), and/or changes in the HER2 protein. Among all molecular subtypes of breast cancer, the HER2 subtype is associated with clinically aggressive outcomes [2]. According to guidelines issued by the American Society of Clinical Oncology and the College of American Pathologists (ASCO/CAP), HER2 testing is mandatory for all patients having invasive breast carcinoma.

Early detection of HER2 is beneficial, as anti-HER2 therapy can be provided to the patient in addition to the regular chemotherapy regimen [3]. The presence of the HER2 subtype also suggests the possibility of a genetic mutation [4]. With the advent of newer Antibody-Drug Conjugate (ADC) therapies that use trastuzumab in HER2-positive breast cancers, the survival rate of breast cancer patients has improved significantly [5]. Immunohistochemistry (IHC) staining is used to determine molecular subtypes of breast cancer, using various antibodies to find overexpression of subtypes in breast tissue, such as Estrogen Receptor (ER), Progesterone Receptor (PR), and HER2. Depending on the amount, intensity, and nature of the brown-colored staining in the IHC assay, histopathologists evaluate different levels of HER2 overexpression. HER2 overexpression level 2+ is equivocal and requires a more advanced diagnosis using the Fluorescent In Situ Hybridization (FISH) technique [6]. Figure 1 shows HER2 overexpression scores based on the IHC staining procedure.

H&E staining and analysis of breast tissue are used to diagnose breast cancer and detect various morphological aspects. These features in H&E-stained tissues also unfold some molecular characteristics expressing subtype information [7]. Deep learning techniques can provide more insights into finding the HER2 subtype directly from H&E-stained images, bypassing the need for an IHC staining procedure for HER2 subtype classification.

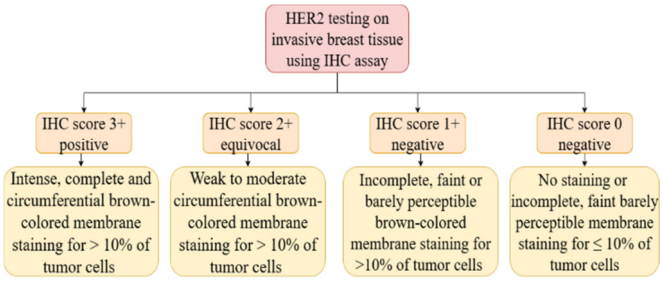

Fig. 1. Evaluation of HER2 overexpression scores using the IHC method.

In many developing countries, there is an insufficient number of pathologists. The IHC assay method is more expensive than H&E staining, and IHC also suffers from intra-observer and inter-observer subjectivity. However, the H&E-staining assay is routinely performed for every breast cancer patient for cancer diagnosis [8]. These factors underscore the importance of using deep learning methods to find molecular subtypes of breast cancer directly from H&E-stained tissue images rather than using the IHC method.

Many pretrained deep-learning-based models have been trained on the globally available ImageNet dataset, facilitating computational pathology research. New deep-learning algorithms are being developed to find molecular subtypes from H&E-stained Whole Slide Images (WSIs) to reduce the burden of manual inspection and analysis. In the last decade, various algorithms have been developed for the diagnosis of breast cancer from histopathological images using deep learning methods and have achieved remarkable results. These artificial intelligence-based systems can help oncologists and histopathologists by serving as digital assistance systems [9]. Deep learning techniques have been proven superior to traditional machine learning techniques for the histopathological analysis of breast cancer. Many studies have used deep learning techniques for the classification of breast cancer. In [10], a model used ResNeXt to extract features, while in [11], a U-Net-based architecture achieved great results in breast cancer tissue segmentation.

II. MATERIALS AND METHODS

This study aimed to determine different levels of HER2 scores directly from H&E-stained WSI patches. Traditionally, IHC-stained WSIs are used for HER2 scoring in breast cancer patients. H&E staining is very cost-effective and is a readily available technique even in remote hospitals.

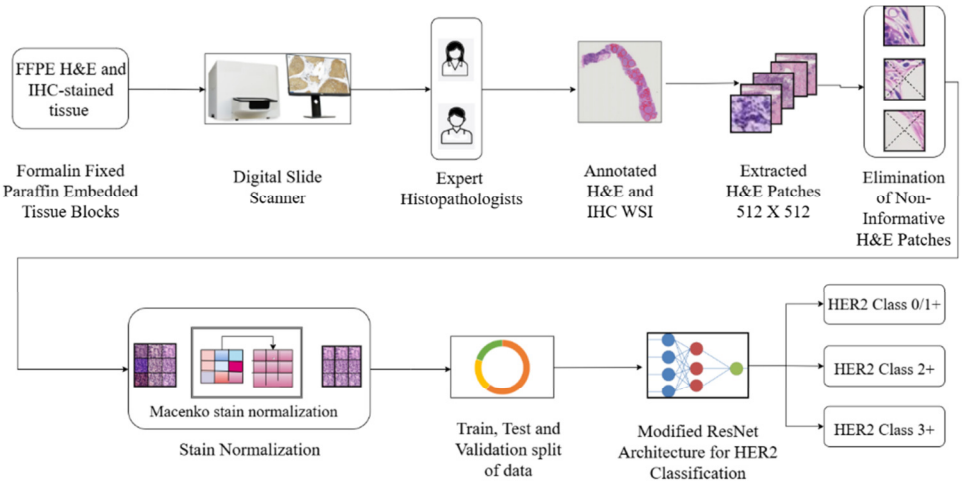

Fig. 2. Architecture of the proposed system.

To find HER2 overexpression from H&E-stained WSIs, a subset of online public data was combined with WSI data collected from a hospital. The dataset consists of H&E-stained WSIs and their corresponding IHC-stained WSIs. Using corresponding IHC-stained WSIs, each H&E-stained WSI was registered in the QuPath software by a suitable angle of rotation. Then, every H&E-stained WSI of the combined dataset was fully annotated for HER2 overexpression using the Region of Interest (RoI) annotation method. For this annotation process, regional correspondence between IHC and H&E images was analyzed. The fully annotated combined dataset was preprocessed, including Macenko stain normalization, patch extraction to sizes of 512×512, and elimination of non-informative patches. The patch extraction process resulted in 3641 patches of label 0/1+, 3357 patches of label 2+, and 3840 patches of label 3+. The total number of patches in the combined dataset is 10,838. The performance of several pre-trained models was tested on the combined dataset by splitting it into train, test, and validation sets. Three pre-trained deep Convolutional Neural Network (CNN) models, namely InceptionV3, EfficientNetB0, and ResNet50, were used as classifiers for the three-class HER2 classification task on the combined dataset. Among these three classifiers, ResNet50 provided the best results in terms of classification accuracy. The accuracy of the ResNet50 base model was improved by inserting Squeeze and Excitation (SE) blocks and an additional depth-wise separable convolutional layer in the baseline architecture. Figure 2 presents the overall architecture of the proposed method. The SE blocks help to incorporate a channel-wise attention mechanism in the ResNet50 architecture. The SE block and the depth-wise separable convolution module are both lightweight architectures that do not cause the parameters of the baseline model to increase significantly. The SE block helps improve feature quality by recalibrating the features corresponding to every channel of the image. Depth-wise separable convolution helps to reduce overfitting of the deep learning model. The best accuracy on the combined dataset was obtained using the modified ResNet50 architecture.

A. Datasets

H&E-stained histopathology images often suffer from stain variations and have complex morphological details. If a histopathology dataset from a single institute is used to train a deep neural network for histopathology image analysis, the model may learn patterns specific to that single cohort. Histopathology datasets from multiple institutes are preferable to ensure the generalized and robust performance of a deep learning model [12]. This study used datasets from two different institutions. The subset of the online Warwick dataset with slide-level HER2 labels was annotated by expert histopathologists from the collaborating hospital for region-level areas. The second dataset was taken from the Symbiosis Medical College for Women (SMCW), Symbiosis University Hospital and Research Center, Pune, India. Multi-institutional data used in this study ensures the validity and stability of the trained models.

1) Warwick Dataset

The University of Warwick has publicly released the HER2 scoring contest dataset to promote artificial intelligence research for HER2 scoring [13]. The Warwick dataset was built as part of the Nottingham HER2 Scoring Challenge. This contest was a man-versus-machine contest intended to automate HER2 scoring directly from H&E-stained WSIs. Ethical approval for the use of patient data was obtained from the Nottingham Research Ethics Committee 2. The contest dataset consists of IHC and corresponding H&E WSIs, gathered from 86 cases of invasive breast cancer. The dataset has ground truths in the form of slide-level HER2 scores and percentage of staining for invasive tumor areas. The ground truth labels are based on the expert opinions of two histopathologists. The histopathology slides of this dataset were scanned and digitized using a Hamamatsu NanoZoomer C9600 scanner. The highest resolution for scanning WSIs was kept at a 40× magnification. Figure 3 shows sample slides of the Warwick dataset. This study utilized a subset of 24 WSIs from the Warwick dataset. An expert histopathologist analyzed the slides, selecting those with a high confidence level in the ground truths provided. Slides of all HER2 scores were chosen to maintain diversity.

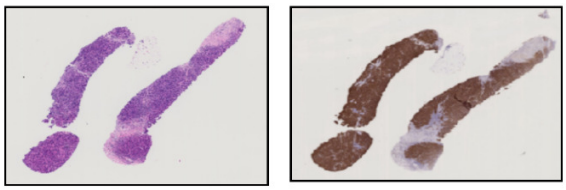

Fig. 3. H&E-stained and corresponding IHC-stained WSIs from the Warwick dataset.

2) Dataset Preparation at the Collaborating Hospital (SMCW Dataset)

The second dataset was collected from Symbiosis Medical College for Women, Symbiosis University Hospital and Research Center (SMCW), after seeking permission from the institute's Independent Ethics Committee (IEC). The patient WSI data was completely anonymized to protect their identity. In addition to IHC and H&E WSIs, no other information was used, such as age, ethnicity, prognosis, etc. A small tissue sample was extracted from the patient's breast using a biopsy procedure. This tissue was preserved and fixed using a formalin solution and oriented using a paraffin wax mold. These Formalin-Fixed Paraffin-Embedded (FFPE) breast tissue blocks were prepared in the collaborating SMCW hospital.

Breast tissue is originally transparent and colorless in nature. To highlight intricate structures in breast tissue, it was stained with H&E dyes. Hematoxylin stains the nuclei blue. The other parts of the breast tissue, such as the intracellular mass, adipose tissue, connective tissue, and the cytoplasm, are pink due to eosin staining. IHC staining is used to find intrinsic molecular subtypes of breast cancer. In the IHC process, to find HER2 overexpression, a monoclonal antibody is used to stain breast tissue. Brown-colored staining is observed in the IHC-stained WSI. The intensity, spread, and appearance of this brown staining depend on the level of HER2 overexpression in a specific slide region. Adjacent tissue sections were cut and prepared for H&E and IHC staining for all patients. These H&E- and IHC-stained WSIs were scanned and digitized using a Morphle 6T Optimus Scanner at the collaborating hospital. This Morphle scanner is capable of scanning 6 whole tissue slides at a time. H&E and IHC slides were scanned using a 40× optical magnification at a resolution of 0.25 µm. The dataset from the hospital consists of 5 WSIs of HER2 scores 0/1+, 4 WSIs of 2+ scores, and 5 WSIs of 3+ scores. This H&E-stained WSI data was annotated by a team of two expert histopathologists from SMCW. The QuPath software was used for the registration of WSIs, and polygonal-shaped annotations were performed on RoIs.

B. Annotations, Patch Extraction, and Data Preprocessing

The subset of the Warwick dataset and the SMCW hospital dataset were combined to form a single dataset with three HER2 classes, namely Zero/One+, Two+, and Three+. These three scores represent the level of HER2 overexpression as negative, equivocal, and positive according to the ASCO/CAP guidelines. This differentiation in HER2 overexpression scores is used to decide treatment strategies for patients and leads to precision medicine. The morphological structure of H&E-stained tissue and IHC-stained tissue is very similar.

Histopathologists annotated H&E-stained WSIs by observing the corresponding IHC-stained WSIs for the adjacent tissue of the same patient. Annotations were performed on the combined dataset. Image registration and annotations were performed using QuPath software version 0.5.1. Image registration of H&E- and IHC-stained WSIs was performed manually by rotating the WSI images in the QuPath software. These annotations were imported into GeoJSON files for H&E and IHC WSIs. These ROI annotations were further used to extract non-overlapping patches of size 512×512 pixels using a script in QuPath software. Figure 4 shows Polygon-shaped ROI annotations.

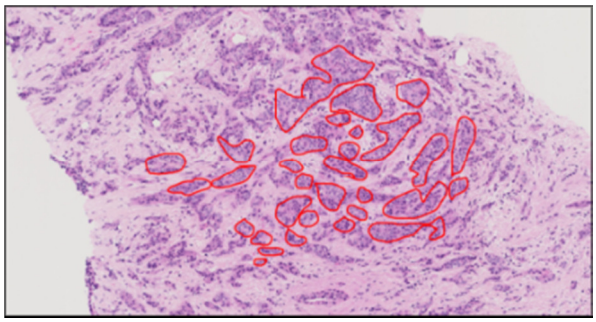

Fig. 4. H&E-stained whole slide images annotated in QuPath.

Usually, WSIs are extremely large, having a resolution of several thousand pixels. Feeding these very high-resolution images to any deep neural network is extremely challenging. Hence, it is common practice in computational histopathology to divide these large WSIs into smaller patches. A review of the literature suggests patch sizes of 224×224, 256×256, 512×512, or 1024×1024 pixels for the supervised learning of AI algorithms. In [14], patches with a 40× zooming magnification factor from online datasets were used for binary detection of breast cancer. In [15], the online Breast Immunohistochemical dataset (BCI) with a patch size of 1024×1024 pixels was used for four-class HER2 overexpression classification. In [16], a weakly supervised deep learning model was used to find HER2 levels based on segmented patch-level data. In [17], a patch size of 512×512 pixels was used for breast cancer classification of histopathology WSIs.

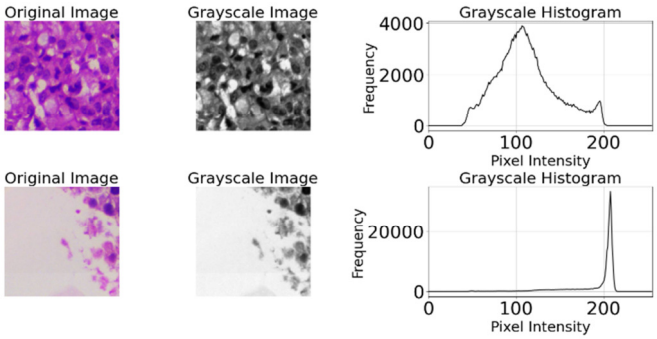

Fig. 5. Difference in grayscale histograms of informative patch (having significant tissue area) and non-informative patch (having less tissue area).

This study used 512×512 patch sizes, as larger dimensions resulted in fewer patches per class. Patch extraction was performed by writing a tiling algorithm in QuPath software. After patch extraction, it was observed that many of the patches consisted of the non-tissue area or the non-informative portion. To exclude such patches, the histograms of such patches were compared with the histograms of normal ones. After analyzing the histograms, all patches consisting of pixel values between 180 and 255 encompassing more than or equal to 60% of the total patch area were excluded from the combined dataset. Figure 5 shows the histograms of an informative patch and a non-informative patch. After filtering out non-informative tiles, the combined dataset consisted of 3539 tiles for class 0/1+, 3077 for class 2+, and 3764 tiles for class 3+.

C. Macenko Stain Normalization

H&E-stained histopathology WSI datasets from more than one institution usually suffer from stain variation problems due to the quantity and quality of dyes used, the shelf life of stains, tissue and reagent handling protocols, etc. The Macenko stain normalization technique was used to solve the issue of stain differences [18]. In this process, all image patches are initially converted to Optical Density (OD) space. The two stains are separated from each other due to this conversion. Stain vectors are extracted using the Singular Value Decomposition (SVD) method. A single reference image from the Warwick dataset was fed as a target image. The stain vectors of all tissue image patches in the dataset were adjusted to match perfectly with the stain vectors of the reference image.

D. Pre-Trained CNNs

Three baseline pre-trained CNN models, Inception V3, EfficientNet B0, and ResNet50, were evaluated for their performance on the combined dataset, each initialized with pre-trained weights from the ImageNet dataset. The bottom layers of the pre-trained network were replaced with custom layers, which contained a global average pooling layer, a fully connected layer, and a final SoftMax function layer for multi-class HER2 classification. Batch normalization was used to help the model converge faster. In this process, the normalization of inputs of a typical layer is performed by scaling the vector of the input data in such a way that the mean of the resultant data is zero and the standard deviation becomes one. This also helps to reduce the inherent covariate shifts of the data.

A dropout of 50% was used to randomly disconnect some neurons in CNN layers to increase generalizability and decrease overfitting. The classification outputs were integers encoded for HER2 overexpression classes. The batch size was kept at 16, selected due to the computational limitations of the available TPU memory. An adaptive learning rate was used, with a starting value of 10^{-4}. The maximum drop in the learning rate was 10^{-7}. To improve model stability and ensure efficient convergence, the learning rate was decreased gradually on the plateau if no improvement in accuracy was found. Table I provides the values of tuned parameters.

TABLE I. TUNED PARAMETER VALUES

Parameter	Optimum value
Batch size	16
Number of epochs	25
Initial learning rate	0.0001
Reduce LROnPlateau by factor	0.5
Verbose	1
Dropout rate	50%
Optimizer	Adam
Total parameters of modified ResNet50	64,412,939
Loss function	Sparse categorical cross-entropy
Hardware execution platform	GPU

E. Modified ResNet50 CNN

An additional SE block was cascaded to the baseline ResNet50 architecture. The inclusion of SE blocks helps improve CNN performance by adding an attention mechanism to the baseline model [19], helping to intensify more useful features by assigning them more weight. The squeeze process applies channel-wise global average pooling on the previous layer. This process helps to squeeze out spatial information at the global-level. This information is excited to get channel-wise attention using dual fully connected layers and activation functions, such as ReLU and sigmoid. Finally, newly learned weights are multiplied by feature maps of earlier network layers to obtain amplified features of important channels. An additional depth-wise separable convolutional block was added at the end of the SE module for further spatial feature refinement. This additional convolutional block ensured an effective spatial representation of the features without adding much computational overhead to the overall architecture. This modification in the architecture of the ResNet50 model enhanced its ability to focus on informative regions.

III. METHOD ANALYSIS

Initially, three pre-trained models were trained and validated on the combined dataset, analyzing their accuracies and confusion matrices on test data. ResNet50 outperformed other classifiers in HER2 classification on patch-level data, achieving an accuracy of 94.03% on the test dataset. Table II shows training and testing accuracy results for different models.

TABLE II. THREE-CLASS HER2 CLASSIFICATION RESULTS BY DIFFERENT CLASSIFIERS AND THE PROPOSED MODIFIED RESNET50 CNN CLASSIFIER.

CNN used for classification	Training accuracy	Testing accuracy
Inception V3	64.73%	64.93%
Efficient Net B0	91.49%	89.34%
ResNet50	96.04%	94.03%
Modified ResNet50	98.70%	98.04%

The modified ResNet50 architecture, enhanced with an SE-block and depth-wise separable convolution, outperformed the three base classifiers and resulted in an accuracy of 98.04%. Figure 6 shows training and validation curves for accuracy and loss parameters. From the graphs, it is evident that, except for some initial spikes, the learning trend of the modified ResNet50 model is more stable and fast converging.

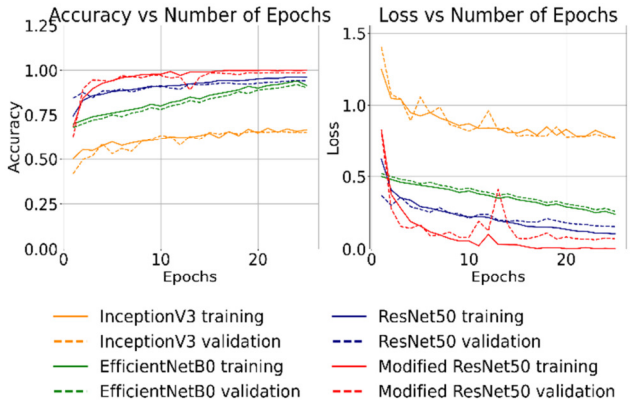

Fig. 6. Training and validation curves for Inception V3, EfficientNet B0, ResNet50, and Modified ResNet50 classifiers.

A. *Implementation Details*

The QuPath coding for patch extraction was implemented in JavaScript. Image preprocessing, stain normalization, model training, and testing were carried out using Python with the help of TensorFlow and SciKit-Learn libraries and Google Colab's v2-8 TPU. The system's RAM utilized was 334 GB. The disk space used was 225 GB.

B. *Confusion Matrices and Metrics*

Figure 7 shows the confusion matrices for three three-class HER2 classifications. In these confusion matrices, rows represent actual labels, while columns indicate predicted labels classified by the respective CNN. The number of correctly classified samples is presented along the diagonals. The number of misclassifications is represented by non-diagonal entries of the matrix.

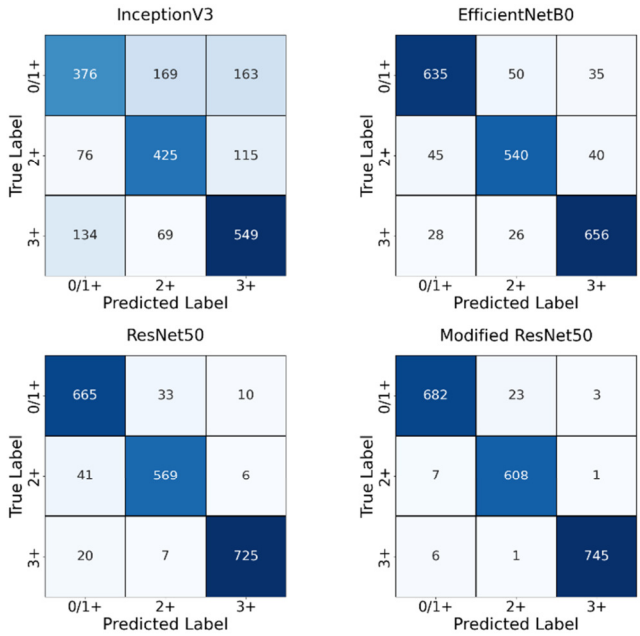

Fig. 7. Confusion matrices for three-class HER2 classification using Inception V3, EfficientNet B0, ResNet50, and Modified ResNet50 classifiers.

C. *Grad-CAM Heatmaps*

Gradient-weighted class activation mapping (Grad-CAM) is used to portray a visual representation of a CNN's classification approach [20], highlighting regions of H&E-stained tissues that influence CNN's final decision-making ability for multi-class HER2 classification. Figure 8 shows a sample for the original heatmap distributions overlaid on the original tissue image patches as indicated.

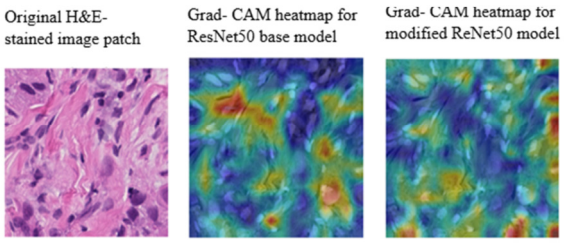

Fig. 8. Grad-CAM heatmap representation of final convolutional layers of ResNet50 CNN and Modified ResNet50 CNN for class 3+.

IV. DISCUSSION AND IMPLICATIONS

This study developed a fully annotated dataset of H&E-stained and IHC-stained slides that represent HER2 overexpression levels. The annotations were performed by observing the regional correspondence between H&E and the supporting IHC WSIs. QuPath was used for image registration and annotations. A part of the Warwick dataset was annotated using the same technique. Both datasets were combined after extracting patches sized 512×512. The non-informative patches in the combined dataset were eliminated by using histogram analysis. The dataset was stain-normalized and preprocessed. Three pre-trained networks were tested on the combined dataset. ResNet-based models have already shown excellence in vision-related tasks. The ResNet50 model was improved by adding a squeeze and excitation block and a depth-wise separable convolution layer, resulting in the best accuracy on this combined dataset. Based on confusion matrices, the modified ResNet50 model reduced the number of misclassifications in terms of false positive and false negative cases. Grad-CAM heatmaps indicated important regions that the model focused on for the final classification layer.

This study used artificial intelligence techniques to manage the diagnostic workflow for breast cancer HER2 classification with better accuracy. Patient stratification becomes easier, as this method provides treatment insights by appropriately classifying the status of HER2 overexpression. This research also helps reduce costs, as preparation costs for H&E-stained WSIs are much lower than those for IHC slide preparation. HER2 identification gives information about the suitability of treatment options and the type of medicines for patients with breast cancer. As a future scope for this study, more annotated images are needed for the real-world breast cancer dataset, which will validate the model's performance on a large dataset. Furthermore, with the help of advanced deep learning algorithms, more insights can be obtained for HER2-low cases with scores 0 and 1+ to find suitable treatment options for patients with HER2-low breast cancer.

V. NOVEL CONTRIBUTIONS AND COMPARISON WITH RELATED WORKS

The research study has the following novel contributions:

- A new dataset of H&E-stained WSIs and corresponding IHC-stained WSIs was created in association with a collaborating hospital (SMCW). This dataset highlights HER2 overexpression levels with RoI annotations and can be used as a test dataset for future studies in computational histopathology.
- The SMCW dataset was combined with a subset of the publicly available Warwick dataset. A total of 10,838 patches of size 512×512 were extracted from the combined dataset. Data from multiple institutions confirm the generalizability and robustness of the model.
- A deep learning-based patch-level HER2 overexpression classification system was designed that can find HER2 status directly from H&E-stained WSIs, bypassing the need for expensive IHC staining for HER2 analysis.
- The accuracy of the baseline ResNet50 model was improved by adding an SE block and depth-wise separable convolution layer.

In [21], WSIs from the Warwick dataset were employed for HER2 overexpression classification using deep learning methods, achieving a patch-level accuracy of 96.17% with the DenseNet201 model. This study also tested a vision transformer model, which outperformed DenseNet201 in terms of other metrics, and obtained a WSI-level accuracy of 96%. In [15], deep learning models were used for HER2 overexpression using patches from the BCI public dataset. Class-wise weight assignment was performed for three deep learning models, and an ensemble model achieved a patch-level HER2 overexpression accuracy of 97.84%. In [22], a customized Inception CNN achieved an accuracy of 85% for patch-level HER2 classification on H&E-stained patches using the BCI public dataset. This study also carried out HER2 classification on IHC images from the same dataset, achieving an accuracy of 88%. In [23], a deep learning classification consisting of three deep CNN models was developed for binary HER2 classification, which achieved an AUC of 82% on the Cancer Genomic Atlas and the Warwick datasets.

VI. CONCLUSIONS

Only a few earlier studies combined WSI datasets from different institutions to form multi-institutional datasets for training deep neural network models. Multi-institutional data help to make the model more generalizable. In addition, there is a lack of fully annotated H&E-stained patch-level data for HER2 classification of breast cancer. This research fused a subset of publicly available Warwick data with data collected from a collaborating hospital. The complete method also encompassed annotating the IHC and H&E-stained WSI data, patch extraction, and preprocessing steps. A modified ResNet50 architecture achieved the best accuracy of 98.04% for three-class HER2 overexpression classifications. The modified ResNet50 model was developed by inserting an SE block and a depth-wise separable convolutional block into the original architecture. The insertion of the SE block resulted in an attention mechanism that provided more weights to useful features. A depth-wise separable convolutional layer cascaded after the SE block helped in enhancing the spatial representation of features. The modified ResNet50 outperformed other models, significantly improving patch-level HER2 classification.

This work demonstrates that the HER2 subtype can be found from H&E-stained WSI patches alone, without the need for tedious and costly IHC staining. Compared to previous studies, the modified ResNet architecture gives higher accuracy for patch-level HER2 classification from H&E-stained tissue images. This study has also developed an annotated paired IHC and H&E WSI dataset. Combining a subset of an online public dataset with a real-world hospital dataset ensures the robustness and generalizability of the model.

REFERENCES

[1] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological Image Analysis: A Review," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 147–171, 2009, https://doi.org/10.1109/rbme.2009.2034865.

[2] D. J. Slamon, G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich, and W. L. McGuire, "Human Breast Cancer: Correlation of Relapse and Survival with Amplification of the HER-2/*neu*Oncogene," *Science*, vol. 235, no. 4785, pp. 177–182, Jan. 1987, https://doi.org/10.1126/science.3798106.

[3] A. C. Wolff *et al.*, "Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer," *Archives of Pathology & Laboratory Medicine*, vol. 147, no. 9, pp. 993–1000, Sep. 2023, https://doi.org/10.5858/arpa.2023-0950-sa.

[4] F. Petrelli, G. Tomasello, S. Barni, V. Lonati, R. Passalacqua, and M. Ghidini, "Clinical and pathological characterization of HER2 mutations in human breast cancer: a systematic review of the literature," *Breast Cancer Research and Treatment*, vol. 166, no. 2, pp. 339–349, Nov. 2017, https://doi.org/10.1007/s10549-017-4419-x.

[5] E. Ferraro, J. Z. Drago, and S. Modi, "Implementing antibody-drug conjugates (ADCs) in HER2-positive breast cancer: state of the art and future directions," *Breast Cancer Research*, vol. 23, no. 1, Dec. 2021, https://doi.org/10.1186/s13058-021-01459-y.

[6] E. A. Rakha *et al.*, "Updated UK Recommendations for HER2 assessment in breast cancer," *Journal of Clinical Pathology*, vol. 68, no. 2, pp. 93–99, Feb. 2015, https://doi.org/10.1136/jclinpath-2014-202571.

[7] N. Naik *et al.*, "Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains," *Nature Communications*, vol. 11, no. 1, Nov. 2020, https://doi.org/10.1038/s41467-020-19334-3.

[8] K. Ingale *et al.*, "Efficient and Generalizable Prediction of Molecular Alterations in Multiple-Cancer Cohorts Using Hematoxylin and Eosin Whole Slide Images," *Modern Pathology*, vol. 38, no. 4, Apr. 2025, Art. no. 100691, https://doi.org/10.1016/j.modpat.2024.100691.

[9] S. A. Joshi, A. M. Bongale, and A. M. Bongale, "Breast Cancer Detection from Histopathology Images using Machine Learning Techniques: A Bibliometric Analysis," *Library Philosophy and Practice (e-journal)*, Art. no. 5376.

[10] R. Gurumoorthy and M. Kamarasan, "Breast Cancer Classification from Histopathological Images using Future Search Optimization Algorithm and Deep Learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12831–12836, Feb. 2024, https://doi.org/10.48084/etasr.6720.

[11] S. M. Shaaban, M. Nawaz, Y. Said, and M. Barr, "An Efficient Breast Cancer Segmentation System based on Deep Learning Techniques," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12415–12422, Dec. 2023, https://doi.org/10.48084/etasr.6518.

[12] M. Tafavvoghi, L. A. Bongo, N. Shvetsov, L. T. R. Busund, and K. Møllersen, "Publicly available datasets of breast histopathology H&E whole-slide images: A scoping review," *Journal of Pathology Informatics*, vol. 15, Dec. 2024, Art. no. 100363, https://doi.org/10.1016/j.jpi.2024.100363.

[13] T. Qaiser *et al.*, "HER2 challenge contest: a detailed assessment of automated HER2 scoring algorithms in whole slide images of breast cancer tissues," *Histopathology*, vol. 72, no. 2, pp. 227–238, Jan. 2018, https://doi.org/10.1111/his.13333.

[14] S. A. Joshi, A. M. Bongale, P. O. Olsson, S. Urolagin, D. Dharrao, and A. Bongale, "Enhanced Pre-Trained Xception Model Transfer Learned for Breast Cancer Detection," *Computation*, vol. 11, no. 3, Mar. 2023, Art. no. 59, https://doi.org/10.3390/computation11030059.

[15] P. G. P, K. Senapati, and A. K. Pandey, "A Novel Decision Level Class-Wise Ensemble Method in Deep Learning for Automatic Multi-Class Classification of HER2 Breast Cancer Hematoxylin-Eosin Images," *IEEE Access*, vol. 12, pp. 46093–46103, 2024, https://doi.org/10.1109/ACCESS.2024.3382212.

[16] P. Jiao *et al.*, "Prediction of HER2 Status Based on Deep Learning in H&E-Stained Histopathology Images of Bladder Cancer," *Biomedicines*, vol. 12, no. 7, Jul. 2024, Art. no. 1583, https://doi.org/10.3390/biomedicines12071583.

[17] T. Araújo *et al.*, "Classification of breast cancer histology images using Convolutional Neural Networks," *PLOS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0177544, https://doi.org/10.1371/journal.pone.0177544.

[18] M. Macenko *et al.*, "A method for normalizing histology slides for quantitative analysis," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Boston, MA, USA, Jun. 2009, pp. 1107–1110, https://doi.org/10.1109/isbi.2009.5193250.

[19] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-Excitation Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 8, pp. 2011–2023, Aug. 2020, https://doi.org/10.1109/tpami.2019.2913372.

[20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, Feb. 2020, https://doi.org/10.1007/s11263-019-01228-7.

[21] S. Kabir *et al.*, "The utility of a deep learning-based approach in Her-2/neu assessment in breast cancer," *Expert Systems with Applications*, vol. 238, Mar. 2024, Art. no. 122051, https://doi.org/10.1016/j.eswa.2023.122051.

[22] M. F. Mridha, Md. K. Morol, Md. A. Ali, and M. S. H. Shovon, "convoHER2: A Deep Neural Network for Multi-Stage Classification of HER2 Breast Cancer," *AIUB Journal of Science and Engineering (AJSE)*, vol. 22, no. 1, pp. 53–81, May 2023, https://doi.org/10.53799/ajse.v22i1.477.

[23] D. Anand *et al.*, "Deep Learning to Estimate Human Epidermal Growth Factor Receptor 2 Status from Hematoxylin and Eosin-Stained Breast Tissue Images," *Journal of Pathology Informatics*, vol. 11, no. 1, Jan. 2020, Art. no. 19, https://doi.org/10.4103/jpi.jpi_10_20.