Breast Cancer Detection Using Histopathology Images Using CNN

Avutu Rahul Reddy
Department of AI
Long Island University
Newyork City, United states
Rahulreddy.avutu@my.liu.edu

Rishik Konduru
Department of AI
Long Island University
Newyork city, United States
Rishik.konduru@my.liu.edu

Shivaram Velidineni
Department of AI
Long Island University
Newyork city, United States
Shivaram.velidineni@my.liu.edu

Shashank Madrri
Department of AI
Long Island University
Newyork city, United states
Shashank.Madrii@my.liu.edu

***Abstract*—Breast cancer is a very dangerous cancer which effects mainly women and young men which can cause havoc in in their lives. Histopathological images are very good at detecting the breast cancer from the microscopic view. But it takes significant time and effort in order to detect the cancer. So, this why CNN is used to detect the cancer in the patients.**

1.INTRODUCTION

The periodical cancer diagnosis remains one of the main causes of mortality due to cancer and tumor diseases among women worldwide. Its early diagnosis has a great impact on the prognosis. In regular clinical practice, histopathological examination under a microscope remains the gold standard for detecting cancerous conditions in the breast. However, this process is extremely tedious and highly dependent on the pathologists' expertise. Intricate tissue architecture and subtle morphological differences between the benign and malignant samples make it much more challenging for consistent and objective diagnoses.

Recent breakthroughs in deep learning have brought computer-aided diagnostic systems into being that can support pathologists by analyzing medical images on their own. Among a number of publicly available datasets on histopathology research, the BreakHis has become a widely used benchmark. The dataset consists of thousands of high-resolution microscopic images of breast tumor tissues, including several benign and malignant subtypes, taken at different magnification levels. This makes BreakHis particularly suitable for training convolutional neural networks to learn discriminative histopathological features.

In the current work, we have applied EfficientNet-B5-a high-capacity convolutional neural network architecture-to perform multi-class breast cancer classification on the BreakHis dataset. EfficientNet models leverage a compound scaling strategy that jointly optimizes network depth and width, together with input resolution, in order to achieve state-of-the-art performance at higher computational efficiency. The B5 variant, in particular, provides higher representational power compared to smaller variants and is thus more suitable for capturing both fine-grained cellular details and global tissue structures in histopathology images.

Advanced training techniques in model robustness and generalization include several incorporations within the learning pipeline. Extensive data augmentation is employed to model real-world variability in the appearance of tissues; besides this, label smoothing and MixUp regularization are adopted as regularizers to help prevent overfitting and improve the calibration of the predictions. EMA of model parameters is a technique used to smoothen the training process and eventually improve results on the validation set. Additionally, test-time augmentation at evaluation time is employed to further enhance the consistency of the predictions. Transfer learning is utilized by fine-tuning an EfficientNet-B5 model pretrained on ImageNet, allowing the network to leverage general visual features and adapt effectively to domain-specific histopathological patterns.

The experimental results show the proposed approach promises to have very good performance on the BreakHis dataset by providing very high accuracy and macro-F1 scores on each of the subtypes in the dataset of breast cancer. Balanced performance on both categories supports the benign and malignant nature, thus showing the effectiveness of the architecture as well as the training strategy. Visual interpretability is added through Grad-CAM, showing regions of interest affecting the model's predictions in a way that can instill clinical trust in the decision-making process of the system. Concluding, this work presents a robust and interpretable deep learning framework featuring the use of EfficientNet-B5 for the classification of breast cancer histopathology. By integrating a powerful CNN architecture with modern regularization techniques and tools for explainability, the proposed method has strong potential to support pathologists and further advance the integration of artificial intelligence into clinical diagnostic workflows.

2. Proposed Approach

This work addresses the problem of multi-class breast cancer histopathology image classification using the BreakHis dataset. The dataset consists of microscopic breast tissue images from eight different pathological subtypes, covering both benign and malignant tumors. A deep learning-based transfer learning approach was adopted to achieve robust and reliable performance in classification due to the visual complexity that characterizes the histopathological images and the class imbalance among tumor subtypes.

The dataset was then divided into training, validation, and testing splits in an appropriate ratio of 70:15:15. Further, it was arranged in a form that can readily be accepted by PyTorch's ImageFolder interface. This structuring makes the work reproducible, as well as label management much easier. Extensive preprocessing and data augmentation were performed on the training set before training, including random resized cropping, horizontal and vertical flipping, rotating, and color jittering. Such augmentations simulate realistic variations in tissue appearance and enhance the generalization capability of the model across previously unseen samples. Validation and test images were preprocessed using only deterministic resizing and normalization for consistency in their evaluation.

In order to alleviate the intrinsic class imbalance within the BreakHis dataset, class-weighted cross-entropy loss was utilized. The frequency of each class was utilized in inverse proportion when calculating the class weights to ensure that all the minority tumor subtypes contributed equally in the loss function during training. Besides, label smoothing was used to avoid overconfidence in the predictions, which helps in improving model calibration-a particularly important aspect of medical image analysis tasks.

At the center of this proposed approach is the convolutional neural network EfficientNet-B5, which is pre-trained on the ImageNet dataset. Among the different members in the EfficientNet family, EfficientNet-B5 offers a better balance between model capacity and computational efficiency by compound scaling of network depth, width, and resolution. To enable transfer learning, the majority of the pre-trained backbone was frozen while only fine-tuning the final convolution block along with the newly introduced classification head. This strategy preserves low-level feature representations while allowing higher-order features to adapt to domain-specific histopathological patterns.

It used the AdamW optimizer, which implements adaptive learning rates further regularized with weight decay for better generalization. A cosine annealing warm restart learning rate scheduler was utilized at the batch level to provide smooth learning rate transitions and avoid abrupt oscillations of the loss. This model also used automatic mixed precision training by default to reduce memory consumption and accelerate training in order to efficiently fine-tune the large EfficientNet-B5 model.

Multiple regularization techniques were integrated into the training pipeline to enhance robustness and reduce overfitting. MixUp augmentation was performed to encourage linear behavior between training samples, which helps improve generalization on limited medical datasets. Gradient clipping was utilized to prevent exploding gradients, ensuring stable optimization throughout training. For additional stability, an EMA of model parameters was maintained and used during validation and checkpointing; this was found to yield smoother and more stable performance compared to raw model weights.

Model performance was evaluated with a variety of metrics, but the most important one was the macro-averaged F1 score. In contrast to simple accuracy, macro-F1 gives equal weight to all classes and hence is much better suited for imbalanced multi-class medical datasets. More in-depth analysis was done by looking at confusion matrices, per-class F1 scores, and detailed classification reports that identified strengths and weaknesses in different tumor subtypes.

At inference time, test-time augmentation was invoked to enhance the reliability of the predictions. In other words, each test image is assessed multiple times under different geometric transformations, and the obtained probability distributions are averaged into the final prediction. This is a kind of ensemble approach that enhances robustness against spatial variations and often gives improved performance without extra retraining of the model. Finally, model interpretability was addressed with the use of Gradient-weighted Class Activation Mapping. Grad-CAM visualizations underlined the discriminative regions within the histopathology images which most strongly influence the model's predictions. These visual explanations provide valuable insight into the decision-making process of the network and help develop trust and transparency, an absolute must for deploying deep learning models in medical applications.

3. ARCHITECTURE

The proposed model is based on the EfficientNet-B5 convolutional neural network architecture, a state-of-the-art deep learning model that is developed for achieving high accuracy with considerable computational efficiency. EfficientNet introduces a compound scaling strategy uniformly to scale network depth, width, and input resolution using a fixed set of scaling coefficients. This balanced scaling approach enables EfficientNet-B5 to capture rich hierarchical features from high-resolution histopathology images without excessive parameter growth, as observed in conventional deep convolutional networks.

EfficientNet-B5 uses a deep convolutional backbone with several stages of MBConv blocks (Mobile Inverted Bottleneck Convolutions). Each MBConv block integrates depthwise

separable convolutions together with squeeze-and-excitation attention mechanisms. Depthwise separable convolutions reduce computational complexity significantly by factorizing the conventional convolution into depthwise and pointwise operations, while the squeeze-and-excitation modules adaptively recalibrate channel-wise feature responses. This allows the network to focus on diagnostically relevant cellular and tissue structures in histopathological images.

It starts with a shallow convolutional stem that processes the input image and extracts low-level spatial features regarding edges and textures. This is followed by a sequence of MBConv blocks, each deeper and larger in channel depth and size of receptive field, respectively. The early layers capture fine-grained visual patterns around cell boundaries or nuclei texture, while deeper layers encode high-order semantic features relative to tissue organization and tumor morphologies. This hierarchical feature extraction is quite critical in histopathology, where both local cellular details and global tissue patterns are relevant to the diagnosis.

In this work, the concept of transfer learning was utilized by initializing EfficientNet-B5 with weights pre-trained on the ImageNet dataset. In order not to overfit on the BreakHis dataset (which is relatively smaller in size), most of the backbone layers were frozen during training to retain general visual representations learned during pretraining. It only unfroze the last convolutional block, which can allow the network to adapt its high-level representations to breast cancer–specific histopathological features while keeping training stability.

The original classification head of EfficientNet-B5 is replaced with a custom classifier tailored to the eight-class BreakHis problem. The modified classifier includes a fully connected layer to reduce the high-dimensional feature vector to a lower-dimensional representation, followed by a ReLU activation to introduce a non-linearity. After that, a dropout layer is applied in order to prevent overfitting by setting the neurons to zero randomly during training. It is then followed by a final fully connected layer, mapping the learned features to eight output neurons, each corresponding to one histopathological tumor subtype. A softmax function is applied implicitly during loss computation to produce class probability distributions. Overall, the EfficientNet-B5 architecture integrates efficient convolution, channel-wise attention, and hierarchical feature extraction into a powerful yet computationally efficient model. Its design is particularly fitted for medical image analysis, given high-resolution inputs are required to capture subtle visual differences between classes, thus imposing a perfect balance of high expressive capacity with robust generalization. The architectural modifications and fine-tuning strategy adopted in this work enable the model to effectively learn discriminative features for multi-class breast cancer histopathology classification while maintaining stability and interpretability.

4.DATASET DESCRIPTION

The Breast Cancer Histopathological Image Classification (BreakHis) is composed of 7,909 microscopic images of breast tumor tissue collected from 82 patients using different magnifying factors (40X, 100X, 200X, and 400X). To date, it contains 2,480 benign and 5,429 malignant samples (700X460 pixels, 3-channel RGB, 8-bit depth in each channel, PNG format)

DATASET consists of :

- → 7909 microscopic images
- → 70% training data
- → 15% validation data
- → 15% testing data

The dataset currently contains four histological distinct types of benign breast tumors: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenona (TA); and four malignant tumors (breast cancer): carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC) and papillary carcinoma (PC).

a model that balances speed, accuracy, and interpretability — all essential qualities for practical clinical deployment.

A. Evaluations

The proposed EfficientNet-B5–based model was evaluated based on multiple performance metrics on a held-out test set, ensuring a reliable and balanced assessment. Considering the class imbalance present in the BreakHis dataset, macro-averaged F1 score was chosen as the primary evaluation metric because it considers all classes equally important. Along with macro-F1, overall accuracy, per-class precision, recall, and F1 scores were calculated. In addition, the model showed very stable training behavior as the training and validation losses smoothly decreased over the course of epochs, which indicates effective optimization and minimal overfitting. The use of early stopping based on validation macro-F1 further ensured optimal generalization performance.

In the test dataset, this model showed a strong classification performance: it showed high accuracy, with a macro-F1 score very high in all eight histopathological classes. Indeed, the confusion matrix demonstrates a well-marked diagonal dominance, meaning that the great majority of the samples were correctly classified. The misclassifications were quite limited and mainly concentrated between histologically similar tumor subtypes, as expected due to subtle morphological similarities presented in some breast tissue samples. Notably, the model guaranteed good balanced performance both for benign and malignant categories, confirming the effectiveness of the class-weighted loss function and of the regularization techniques adopted during the training process.

This was further supported by the qualitative evaluation using the Grad-CAM technique, which gave more insights into the model's decision-making process. The visual explanations showed that clinically relevant areas such as cellular nuclei,

glandular structures, and abnormal tissue patterns are the focuses of the network in making predictions. This agreement between model attention and meaningful histopathological features enhances interpretability and increases the trustworthiness of the system. As a whole, these results of evaluation validate our proposed approach as being robust, generalize well to unseen data, and suitable for multi-class classification tasks of breast cancer histopathology.

B. Conclusion & future work

In this work, a deep learning framework for multi-class classification of breast cancer histopathology images based on the EfficientNet-B5 model was introduced. Based on transfer learning from ImageNet and incorporating sophisticated regularization methods like data augmentation, MixUp, label smoothing, and Exponential Moving Average, it was shown that state-of-the-art performance can be achieved on the BreakHis dataset. Use of the macro-F1 measure as a main performance metric ensured that equal importance is given to all classes, including minority classes. It was shown that the classification model is capable of identifying discriminative features and generalizing well on new images. Also, exploration with Grad-CAM visualization added more transparency and relevance to the proposed method.

Despite these encouraging findings, several avenues for future work still exist. First, patient-level data splitting might be employed to better mitigate risks associated with data leakage. Second, exploring methods for extending this approach that combine multi-magnification information or attention fusion techniques could potentially enhance classification accuracy. Third, testing with different architectures, perhaps involving Vision Transformers or a CNN-Transformer combination, might offer some additional benefits. Finally, testing on larger multicenter datasets and developing an implementation strategy within a clinical decision-support platform would be an exciting area for additional research and testing.

ACKNOWLEDGMENT

I would like to take this opportunity to show my appreciation to my project supervisor and team members for continuous guidance, encouragement, and providing valuable feedback in relation to this project. I am also grateful to researchers and organizations who provided openly available datasets of medical images, such as BreakHis, which were useful in training and testing the model. Their contributions greatly facilitated research into medical image analysis

REFERENCES

[1] Deep learning-based risk stratification of preoperative breast biopsies using digital whole slide images (2024) — Boissin, C., Wang, Y., Sharma, A. et al.

[2] Extended Deep-Learning Network for Histopathological Image-Based Multiclass Breast Cancer Classification Using Residual Features (2024) — Mewada, H.

[3] Deep learning and vision transformers-based framework for breast cancer and subtype identification (2025) — Author(s) not specified here, published in *Neural Computing and Applications*.

[4] Breast cancer diagnosis with MFF-HistoNet: a multi-modal feature fusion network integrating CNNs and quantum tensor networks (2025) — Mahmood, T., Saba, T., Rehman, A.

[5] "Treatment episode data set: discharges (TEDS-D): concatenated, 2006 to 2009." U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Office of Applied Studies, August, 2013, DOI:10.3886/ICPSR30122.v2

[6] Comparative Analysis and Ensemble Enhancement of Leading CNN Architectures for Breast Cancer Classification (2024) — Murphy, G., Singh, R. (pre-print on arXiv)

[7] Enhanced Multi-Class Breast Cancer Classification from Whole-Slide Histopathology Images Using a Proposed Deep Learning ModelAdnan Rafiq 1, Arfan Jaffar 1, Ghazanfar Latif 2, 3,Sohail Masood 1 and Sherif E. Abdelhamid 4