Enhancing Breast Tumor Classification in Microscopic Images through Fourier Transform-Based High Pass Filtering

Serhii Biletskyi
dept. of Sofware Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
serhii.biletskyi1@nure.ua

Volodymyr Kobziev
dept. of Sofware Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
volodymyr.kobziev@nure.ua

***Abstract*—Breast cancer remains a significant health concern globally, and early detection plays a pivotal role in improving patient outcomes. With the advent of digital pathology, computer vision techniques have emerged as powerful tools for automated analysis of microscopic breast tumor tissue images. The efficacy of Fourier Transform-based high pass filtering in enhancing the classification performance of a Convolutional Neural Network (CNN) model trained on microscopic breast tumor tissue images is investigated in this study. Initially, attempts to train the model on grayscale images yielded unsatisfactory results, indicating the limitations of conventional image representation for this task. However, notable improvements in classification accuracy were observed through the application of high pass filtering with varying radii, particularly with smaller radii. Furthermore, the analysis revealed a correlation between the radius of thehigh pass filter and model generalization, with larger radii leading to increased overfitting. These findings underscore the potential of Fourier Transform-based preprocessing techniques in augmenting the discriminatory power of CNN models.**

Keywords—Fourier Transform, CNN, image preprocessing, classification, breast cancer, microscopic images

I. Introduction

Breast cancer is one of the most prevalent malignancies affecting women worldwide, with early detection being cru- cial for successful treatment outcomes. Traditional diagnostic methods rely heavily on manual examination of histopatho- logical images by trained pathologists, a process that can be time-consuming and prone to interobserver variability. The advent of digital pathology has paved the way for computer- aided diagnosis systems, leveraging advanced image analysis techniques to assist pathologists in accurate and efficient diagnosis.

Despite the promise of computer vision in automated anal- ysis of histopathological images, challenges persist in devel- oping robust models capable of accurately distinguishing be- tween benign and malignant breast tumor tissue. Conventionalapproaches often rely on raw image representations whichmay not fully capture the discriminative features present in microscopic images.

The primary objective of this study is to investigate the efficacy of Fourier Transform-based high pass filtering in enhancing the classification performance of Convolutional Neural Network (CNN) model [1, 2] for breast tumor tissueimages. Specifically, we aim to:

- Evaluate the performance of CNN models trained on grayscale images for breast tumor classification.
- Explore the impact of Fourier Transform-based high pass filtering on image representation and classification accuracy.
- Investigate the relationship between the radius of the highpass filter and model generalization.

II. Literature Review

One notable approach in the computer-aided diagnosis (CAD) of breast cancer is the use of Convolutional Neural Networks (CNNs) for image classification. Studies [3–5] have demonstrated the effectiveness of CNNs in distinguishingbetween benign and malignant tissue. These models leverage the rich hierarchical features learned directly from the image data, thereby enabling automated classification with high ac- curacy.

Image preprocessing plays a crucial role in enhancing the performance of CAD systems for breast cancer diagnosis. Previous work, such as [6] has proposed image normalization methods to improve the consistency and quality of histopatho- logical images for quantitative analysis. These methods aim toreduce variations in image appearance due to factors such as staining variability and tissue thickness, thereby improving the robustness of classification models.

In the context of this thesis, the focus shifts towards ex- ploring alternative preprocessing methods, specifically FourierTransform-based high pass filtering [7]. Fourier Transform- based techniques have been widely used in image processing for enhancing image contrast and edge detection. By applying high pass filtering with varying radii, this approach aims to highlight relevant features in histopathological images, thereby improving the discriminative power of classification models.

III. Methodology and Research

A. Dataset Description

The Breast Cancer Histopathological Image Classification (BreakHis) dataset [8] serves as the foundation for this study. It comprises 9,109 microscopic images of breast tumor tissue obtained from 82 patients, captured at varying magnification levels (40X, 100X, 200X, and 400X). The dataset contains 2,480 benign and 5,429 malignant samples, each with dimen- sions of

700x460 pixels and encoded in 3-channel RGB formatwith 8-bit depth per channel.

B. Preprocessing

Prior to model training, several preprocessing steps were ap-plied to the BreakHis dataset. Firstly, all images were resized to a standardized dimension of 350x230 pixels. Subsequently, the dataset was partitioned into train, validation , and test sets, with 80%, 10%, and 10% of the images allocated to each subset respectively. To address class imbalance, the minority class (benign samples) within the training set was oversampled to achieve a balanced distribution of classes, specifically targeting a 50:50 ratio between benign and malignant samples. Additionally, as part of the experimental design, images were converted to grayscale to facilitate a comparison between models trained on raw grayscale images and those augmented with Fourier Transform-based high pass filtering.

C. Model Training

For the binary classification task, a simple CNN architecture (see Figure 1) was employed due to its effectiveness in image classification tasks. The model was trained using binary crossentropy loss and the Adam optimizer with a learning rate of 0.001. Each training session consisted of 10 epochs, with performance metrics including accuracy, precision, recall, F1 score, and AUC-ROC monitored throughout training.

Each training session was repeated 3 times. Consequently, the reported results, including metrics observed during train- ing, represent the means calculated across these 3 sessions.

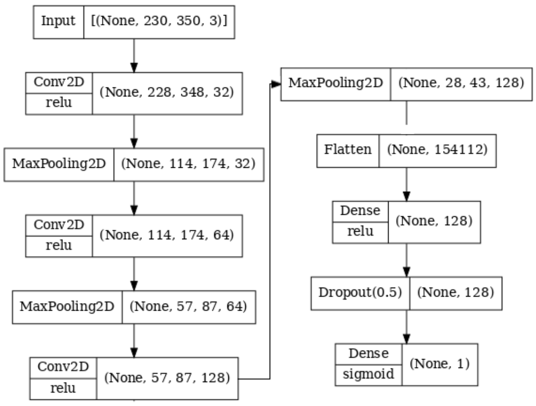

Fig. 1. The architecture of the CNN model for binary classification task.

D. Fourier Transform-based High Pass Filtering

After applying Fourier Transform-based high pass filtering using ideal circle masks with radii of 1, 3, 5, 10, 20, and40 to augment the grayscale images, the transformed image is represented in the frequency domain as complex numbers. When this transformed image is brought back to the spatial domain, we retrieve complex numbers. However, in this study,only the real parts of these complex numbers are utilized, resulting in a final image representation with floating-point pixel values. Notably, due to the nature of the filtering process, some of these pixel values may be negative, which is an inherent characteristic of the filtering technique and should be taken into consideration during subsequent analysis or visualization (see Figure 2). This transformation and filtering strategy aim to accentuate fine details and edges, thereby potentially aiding in the discrimination between benign and malignant tissue in medical image analysis applications.

IV. Experimental Results and Analysis

A. Model Performance without Fourier Transform

Training the CNN model on raw grayscale images yielded stagnant results, as evidenced by the constant AUC-ROC value of 0.5 throughout training and validation (see Figure 3). This suggests that the model failed to effectively learn discrimi- native features from the grayscale representation, resulting in random guessing for classification.

B. Effect of Fourier Transform-based High Pass Filtering

Upon applying Fourier Transform-based high pass filtering with varying radii, notable improvements in classification performance were observed. Models augmented with smaller radii, such as 1, 3, and 5, achieved higher AUC-ROC val- ues at the end of training and validation, reaching up to

0.76. However, as the radius increased, the AUC-ROC values exhibited a decreasing trend. For instance, the model witha radius of 40 achieved a lower AUC-ROC of 0.62. This trend indicates that smaller radii preserve more discriminative information, while larger radii lead to information loss and decreased classification performance.

C. Comparison of Results Across Models

Comparative analysis revealed significant differences in performance between models trained on raw grayscale images and those augmented with Fourier Transform-based high pass filtering. The latter demonstrated much better classification ability, particularly for smaller radii, highlighting the impor- tance of frequency domain information in enhancing feature representation for breast tumor classification.

Analysis of the test metrics (see Figure 4) revealed a clear correlation between the radius of the high pass filter

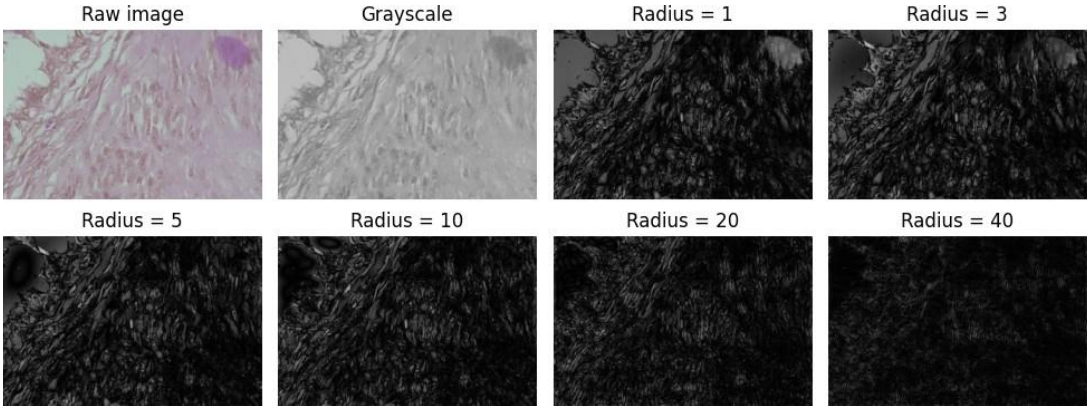

Fig. 2. Example of applying high-pass image filtering with different mask radius values; for visualization, absolute pixel values were taken and multipliedby 3.

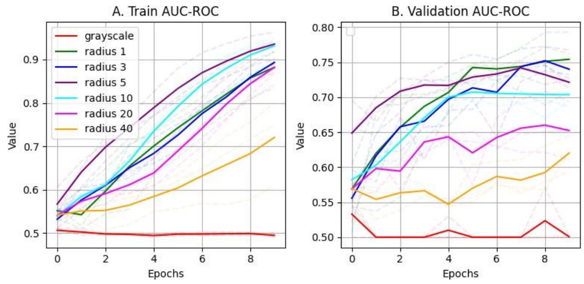

Fig. 3. The graph illustrates the AUC-ROC metric values over the course of training for grayscale and filtered images at varying mask radii, for both the training dataset (A) and the validation dataset (B).

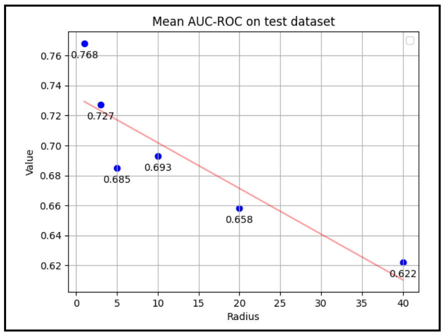

Fig. 4. The graph illustrates the correlation between the AUC-ROC metricand the values of the high-pass filter mask radius on the test dataset

and the mean AUC-ROC value. Models augmented with smaller radii exhibited higher mean AUC-ROC values, with radius 1 achiev-ing a mean AUC-ROC of 0.768. Conversely, as the radius increased, there was a corresponding decrease in the mean AUC-ROC value, with radius 40 yielding a mean AUC-ROC of 0.622. This further supports the notion that smaller radii are more effective in preserving discriminative information within the images, while larger radii lead to information loss and decreased classification performance.

V. Conclusion and Future Investigation

In conclusion, this thesis has explored the effectiveness of Fourier Transform-based high pass filtering in enhancing breast tumor classification using microscopic images. By ap- plying ideal circle masks with varying radii, the study demon- strated significant improvements in classification accuracy on grayscale images, particularly with smaller radii.

Looking ahead, future research should not only focus on refining existing methodologies but also explore novel ap- proaches to further enhance breast tumor classification. One promising avenue is to investigate alternative variations of Fourier Transform-based filters or other frequency domaintransformations. Additionally, exploring the integration of complementary information, such as phase information or directional filters, could offer new insights into improvingclassification performance.

References

[1] A. Krizhevsky, I. Sutskever and G. Hinton, “ImageNet classification with deep convolutional neural networks.” Communications of the ACM, vol.60, pp. 84-90, May 2017, doi:10.1145/3065386.

[2] Khovrat, A., Kobziev, V. Using Recurrent and Convulation Neural Networks to Indentify the Fake Audio Messages. In: *2023 IEEE 7th International Conference on Methods and Systems of Navigation and Motion Control, MSNMC 2023 - Proceedings*, 2023, pp. 174–177.

[3] G. Litjens et al., “Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis.” Scientific Reports, vol. 6, 2016, doi:10.1038/srep26286.

[4] K. Sirinukunwattana et al., “Locality Sensitive Deep Learning for De-tection and Classification of Nuclei in Routine Colon Cancer Histology Images.” IEEE transactions on medical imaging, vol. 35, pp. 1196-1206,2016, doi:10.1109/TMI.2016.2525803.

[5] T. Araújo et al., “Classification of breast cancer histology images using Convolutional Neural Networks.” PloS one, vol. 12, 1 Jun. 2017, doi:10.1371/journal.pone.0177544.

[6] M. Macenko et al., ”A method for normalizing histology slides for quan- titative analysis,” 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, Boston, MA, USA, 2009, pp. 1107-1110,doi: 10.1109/ISBI.2009.5193250.

[7] R. Gonzalez, Z. Faisal, Digital Image Processing Second Edition, 2016.

[8] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, ”A Dataset for Breast Cancer Histopathological Image Classification” in IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1455-1462, July 2016, doi: 10.1109/TBME.2015.2496264.