Comprehensive Experiments on Breast Cancer Hematoxylin and Eosin-stained Images using UNet

Emily Jackson
Kennesaw State University
Marrietta, Georgia, USA
ejack150@students.kennesaw.edu

Faye Le
Kennesaw State University
Marrietta, Georgia, USA
nle30@students.kennesaw.edu

Je'Dae I. Lisbon
Kennesaw State University
Marrietta, Georgia, USA
jlisbon1@students.kennesaw.edu

Max Coleman
Kennesaw State University
Marrietta, Georgia, USA
jcole186@students.kennesaw.edu

Jordyn Burman
Kennesaw State University
Marrietta, Georgia, USA
nburman@students.kennesaw.edu

Astrid Wonderley
Kennesaw State University
Marrietta, Georgia, USA
awonderl@students.kennesaw.edu

Sepehr Eshaghian
Kennesaw State University
Marrietta, Georgia, USA
seshaghi@students.kennesaw.edu

Sanghoon Lee
Kennesaw State University
Marrietta, Georgia, USA
slee297@kennesaw.edu

ABSTRACT

The current challenge in breast cancer segmentation on hematoxylin and eosin-stained images lies in accurately capturing histologic phenotypes associated with cancer biomarkers. Many researchers have studied histologic patterns of cancerous regions of breast cancer images for a better understanding of diagnosis and treatment, yet it is not fully investigated because of its variability, complexity, and large data volumes. Comprehensive experiments on breast cancer segmentation can address this challenge by identifying heterogeneous cell regions in the tumor microenvironment and investigating a new methodology of segmenting histologic images by using advanced deep learning architectures. In this paper, we present findings in three experiments on breast cancer segmentation, exploring the effectiveness of a convolutional neural network called UNet identifying multiple heterogeneous cells such as blood, blood vessels, fat, glandular secretions, necrosis, and plasma cells in a tumor microenvironment, investigating the performance of UNet architecture in different sizes of cancerous regions such as a tumor, and tumor-infiltrating lymphocytes, and proposing a new methodology of a neural-style guided data augmentation focusing on the image segmentation of breast tumor-related nuclei in hematoxylin and eosin-stained images. The experiment results show that a modified UNet performs well on fat identification. Another experiment shows that the smaller input size of the tumor, stroma, and tumor-infiltrating lymphocyte provides better performance than the larger input size. Moreover, we demonstrate that the proposed method outperforms the traditional deep neural network models.

ACMSE 2024, April 18–20, 2024, Marietta, GA, USA

ACM ISBN 979-8-4007-0237-2/24/04.
https://doi.org/10.1145/3603287.3651207

CCS CONCEPTS

• **Computing methodologies** → **Image segmentation**.

KEYWORDS

Breast Cancer, Convolutional Neural Networks, Image Segmentation, Data Augmentation

ACM Reference Format:
Emily Jackson, Faye Le, Je'Dae I. Lisbon, Max Coleman, Jordyn Burman, Astrid Wonderley, Sepehr Eshaghian, and Sanghoon Lee. 2024. Comprehensive Experiments on Breast Cancer Hematoxylin and Eosin-stained Images using UNet. In *2024 ACM Southeast Conference (ACMSE 2024), April 18–20, 2024, Marietta, GA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3603287.3651207

1 INTRODUCTION

Recent advances in digital pathology have allowed pathologists to use computer machines to assist them in better diagnosis of cancer, improving the diagnostic process for patients [25, 37]. As computer vision and deep learning are revolutionizing cancer-related medical image analysis assisting early cancer detection for improving patient outcomes, it is not evitable for cancer researchers to utilize deep learning algorithms for identifying characteristics of human tissue cells to determine a patient's health problems [9, 33].

Breast cancer is one of the most significant global health challenges, with early detection being pivotal for improving patient outcomes and standing as the foremost cause of cancer-related mortality among women worldwide [38]. With 287,850 new cases reported in 2022 [4], breast cancer maintains its status as the most prevalent cancer. Various diagnostic tools have emerged to generate visual and functional representations of the human body and organs for non-invasive clinical analysis through X-ray-based techniques such as conventional X-ray [14, 26], computed tomography (CT) [10, 39], and mammography, as well as molecular imaging, magnetic resonance imaging (MRI), and ultrasound (US) imaging [12]. These imaging-based analyses often initiate breast cancer diagnosis, followed by a crucial histopathological analysis of breast

tissue samples. This latter process necessitates precise cell nucleus identification and classification within tumors, playing a pivotal role in determining tumor characteristics such as grade, subtype, and aggressiveness. The accurate identification of nuclei is instrumental not only in diagnosis but also in the effective planning of treatment strategies [7, 17, 22, 27].

While X-ray-based techniques have improved rates of early detection substantially, it is markedly less effective for women with dense breasts, who already suffer from an increased risk of breast cancer later in life [36]. Computer vision, and specifically deep-learning models, can assist radiologists in identifying cases of breast tumor identification and classification by reaffirming their conclusion, and with enough improvement, eventually able to identify breast cancer unassisted. Breast cancer is a particularly difficult condition to diagnose, as the prognostic parameter changes with the progression of the cancer. For the scope of this paper, refinement of state-of-the-art deep learning modalities is crucial in adapting widely used, successful methods, such as UNet, VGG, and ResNet series, to some of the more granular challenges specific to breast cancer, in addition to improving the performance of models on diagnosing based on medical images.

Medical image analysis is a crucial component of modern healthcare, enabling precise diagnoses and guiding treatment decisions [34]. Traditional deep learning models in medical image analysis use a set of fixed-sized images and output a predicted label for each pixel along with a confidence status. UNet is one of the representative convolutional-wise deep-learning models primarily used in medical image analysis [32]. The emergence of the UNet architecture has opened new avenues in automating and refining the process of specific cell identification and analysis in histopathological images, but there is still a need to improve the quality of semantic image segmentation in medical image analysis. Since the UNet architecture is characterized by a U-shaped structure capturing context information through the encoder-decoder process, UNet has been combined with various deep learning models such as ResNet [13], VGG [24], and Generative Adversarial Networks (GANs) [5]. This paper will introduce a groundbreaking method that combines the UNet architecture with VGG16 to enhance medical image analysis, particularly focusing on the segmentation of breast tumor cell nuclei. This integration is expected to revolutionize the accuracy and efficiency of histopathological image analysis, which is essential for identifying and delineating structures within these images.

While the integration of deep learning models in medical imaging, particularly for breast cancer diagnosis, offers significant advancements, it is imperative to acknowledge inherent drawbacks and limitations such as unit-emphasized experiments merely focusing on phenotypes of specific tissue cells fixing a size of images in advance. Deep-learning models used for identifying specific types of cancer cells have presented impressive capabilities for automatic cancer-related image predictions in tumor microenvironments [23, 40]. However, they have limited use in terms of identifying the effect of their models with different input sizes which remains a challenging issue to overcome. Comprehensive experiments on deep learning models with different input sizes are necessary for better identification of cancer cells.

In this paper, we propose a new deep learning method for integrating UNet with the VGG16 architecture using a neural-style guided data augmentation focusing on the image segmentation of breast tumor-related nuclei in hematoxylin and eosin-stained images, leveraging the integrated approach to improve the accuracy and efficiency of nuclei segmentation in breast cancer, thereby aiding in early and precise cancer diagnosis. The proposed method involves two key strategies: data augmentation and the fusion of two distinct neural network architectures. The data augmentation approach employs a Neural Style Network (NSN), an innovative technique that transfers artistic styles to medical images. This method aims to diversify the dataset, thereby enhancing the model's capacity to generalize across various styles and complexities encountered in histopathological images. At the heart of this method is the fusion of the UNet architecture, known for its good performance in semantic image segmentation, with the VGG16 model, which is celebrated for its robust feature extraction capabilities. The proposed model combines the pixel-level accuracy and detailed segmentation capabilities of UNet with the feature learning of VGG16. This fusion is anticipated to significantly improve accuracy and precision in medical image classification. Moreover, we explore the effectiveness of UNet by identifying six heterogeneous cells: blood, blood vessels, fat, glandular secretions, necrosis, and plasma cells through comprehensive experiments. We also investigate the performance of UNet architecture in different sizes of three cancerous regions: tumor, and tumor-infiltrating lymphocytes.

The rest of the sections in this paper are organized as follows. In Section 2, we describe the related works on medical image segmentation in the field of biomedical research. Section 3 details the proposed VGG16 encoded UNet architecture with a neural style transfer method. Section 4 presents the comprehensive experiments on breast cancer image segmentation on hematocylin-eosin stained images. The conclusion is given in Section 5.

2 RELATED WORK

2.1 Medical Image Analysis

Medical image analysis has been broadly studied in the field of biomedical research by providing meaningful information from digitized images obtained through different medical imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), and whole slide imaging (WSI). Baracos et al. [3] presented a population-based experiment for studying body composition in cancer patients using diagnostic computed tomography images. Holli et al. [15] studied texture information extracted from breast magnetic resonance images to distinguish between normal breast tissue and invasive breast cancer. Cruz-Roa et al. [6] presented an invasive breast cancer detection method on whole slide images by proposing a convolutional neural network model. The qualitative or quantitative information extracted from the digitized images through various computational models can now provide valuable insights into the interesting regions of images enabling clinicians to detect specific diseases within the human body.

2.2 UNet and Variants

The evolution of computational models plays a crucial role in advancing diagnostic accuracy and treatment efficacy [8, 20]. Since

UNet has received much attention in the biomedical research field as a primary use in medical image analysis, the UNet architecture has been modified for medical image segmentation [35]. The attention UNet is one of the variants of the basic UNet presented for capturing and handling fine details in medical images [29]. This variant also has received a lot of attention because of its attention mechanisms allowing the neural network to focus on related regions of the input image to enhance the ability of the model. Some studies have been done by combining the residual concept of skipping connections with the basic UNet architecture [1, 42]. They incorporated the residual connections to address the vanishing gradient problem and showed the effectiveness of the variant of the UNet architecture.

2.3 UNet and VGG16

UNet can be directly integrated with VGG-16 as an encoder model to improve the image segmentation of brain tumors [11]. The integrated model was introduced to identify regions of tumor cells, showing the achievement with pixel accuracy of 99.75% with the integrated UNet architectures. The significant improvement in accuracy over common CNN-based architectures motivated us to use VGG16 as an encoder model to enhance the effectiveness of cancerous region image segmentation. Anindya et. al. presented the application of a UNet-VGG16 model, supplemented with transfer learning, which demonstrates a high Correct Classification Rate (CCR) in identifying brain tumor areas [30]. Their study achieved a mean CCR value of about 95.69%, highlighting the model's efficiency in recognizing tumor regions in MRI scans. Their paper also motivated us to design a CNN-based architecture through the combination of UNet and VGG16 architectures for accurate medical image analysis. Another study of UNet combined with VGG16 was presented for further enrichment with transfer learning and dropout regularization [31]. The main objective was to achieve clear tumor visualization in image segmentation. This hybrid architecture demonstrates the potential of UNet and VGG16 in deep learning for precise image segmentation, particularly in medical imaging applications.

2.4 UNet and Augmentation

While the integration of UNet and VGG-16 demonstrates the effectiveness of image segmentation, several image augmentation methods have been utilized in medical image analysis. In particular, Yin et al. [41] proposed a context-aware generative method altering the grayscale of CT scans with minimal semantic loss to enhance the performance of an existing segmentation model. They followed a traditional style transformation pipeline for augmenting images and showed an improvement of between 2% and 4UNet% in pixel segmentation accuracy over the original UNet for spine segmentation. Their study emphasizes the importance of preserving semantic integrity in medical image processing and demonstrates how style transfer can be effectively employed for data augmentation.

The exploration of previous studies elucidates the significant strides made in the realm of medical image analysis. The integration of UNet and VGG16 architectures, as demonstrated in various studies, underscores their efficacy in precise image segmentation, particularly in the context of medical image analysis [11, 18]. Furthermore, the adoption of neural-style networks for data augmentation marks a novel approach to enhancing model generalization across diverse histopathological images [16, 19]. Collectively, these studies provide a robust foundation and invaluable insights for the current research, emphasizing the continual innovation required in the quest for improved healthcare outcomes through advanced medical imaging techniques. We will focus on the integration of UNet with VGG16 for enhanced image segmentation and the innovative use of Neural Style Networks in data augmentation. Moreover, we will perform comprehensive experiments on histopathology images from whole slide imaging.

3 COMPREHENSIVE EXPERIMENTS ON CANCER-RELATED IMAGE SEGMENTATION USING UNET

3.1 Neural Style-equipped UNet-VGG16 Model

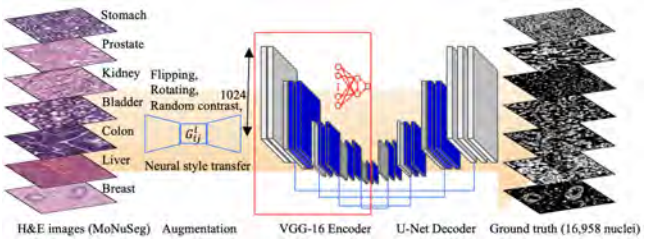

Figure 1: The Overall Process of the Neural Style-equipped UNet-VGG16 Model

3.1.1 Method. The basic idea of the proposed method is to utilize the combination of the VGG16 encoder and the UNet decoder. The model experienced a bottleneck at the point where it had 1024 channels. This concatenate layer was the bridge that connected the encoder of VGG16 to the decoder of UNet. We took the output of the last convolutional layer of each shape (512 x 512 x 64, 256 x 256 x 128, 128 x 128 x 256, 64 x 64 x 512) in the encoder and passed it into the decoder as input. This skip connection reduced the information lost during the encoding process and helped reconstruct the high-resolution information from the earlier layers. We then modified the hybrid model by adding more convolutional layers along with dropout layers. We increased the number of feature maps to 1024. The previous encoder only had 512 feature maps. The new model experiences a bottleneck at the bridge where it has 2048 channels. Although we still passed the output of the last convolutional layer of each shape into the decoder, the number of input shapes increased because of the extra convolutional layers. The addition of the input shape 32 x 32 x 1024 helped improve the model's ability to capture intricate patterns and context. Moreover, the introduction of dropout layers decreased the overfitting problem the previous model was having.

To increase the number of training samples in this paper, we implemented several image augmentation techniques such as flipping the image and mask left or right, rotating the image and mask random increments of 90 degrees (maximum 270), randomizing

the contrast of the image, and applying neural style transfer to the image with Van Gogh's *The Starry Night*. For the neural style transfer, we chose layers from a pre-trained VGG16, with frozen ImageNet weights, to serve as the *content layers* and *style layers* in our style transfer model.

Specifically, we used all of the input layers from VGG16, the outputs from *block5-conv2* to represent the content layer, and the *block1-conv1*, *block2-conv1*, *block3-conv1*, *block4-conv1*, *block5-conv1*, outputs to represent the style layers. For the training of the style transfer model, first calculate the style of the image, which is typically done by calculating the Gram matrix:

$$G^l_{ij} = \sum_k F^l_{ik} F^l_{ijk} \tag{1}$$

where l is the *style layer* the current gram matrix is being calculated for, i and j are feature maps from the output of the current *style layer*, and k is the number of channels. The Gram matrix is the dot product of each feature map in the layer with every other feature map in that layer. We used the calculated style of the images and the output from the single *content layer* to calculate style loss, the mean squared error between the style of the training image and the style of *The Starry Night*, and content loss, the mean squared error between the training image from the previous neural style transfer iteration, and the current one. We finally added the style loss and content loss together to give us a total loss and performed a gradient descent to minimize this total loss. The overall process of the proposed method is shown in Figure 1. H&E images are preprocessed, the augmented images are created, and the input images are fed to the VGG-16 encoder. The decoder of the UNet is used to predict nuclei regions. We used Adam Optimizer. For the loss function, we decided to use Binary Cross Entropy (BCE) at first. BCE is suited for semantic segmentation where the goal is to classify whether a pixel belongs in the foreground or not. However, after seeing the results, we realized there was a class imbalance in the dataset. To address this problem, we implemented Binary Focal Cross Entropy (BFCE). BFCE introduced a modulating factor to down-weight the contributions of well-classified samples. The model was trained on augmented data for better generalization.

3.1.2 Experiment Results. In this section, we experiment to verify the effectiveness of the proposed method by comparing it with the baseline of the model and the state-of-the-art model called UNet++ EfficientNet.

We used a refined version of the dataset from Dinh et al., which consists of a combination of images from the MonuSeg-2018 dataset [21] and the Triple Negative Breast Cancer (TNBC) dataset [28]. MonuSeg-2018 was published as part of a semantic segmentation challenge at the MICCAI 2018 conference in Spain and contains hematoxylin and eosin (H&E) stained medical images of tumors in various organs. On the other hand, the TNBC dataset was created at Curie University in Paris for P. Taylor et al. and contains H&E stained images of breast cancer. Our dataset consists of 31 training images, all from MonuSeg-2018, that are augmented up to 100 training images, while our test folder contains 50 images from the TNBC dataset. The evaluation metrics used in this paper include Intersection over Union (IoU), F1-score (F1), and accuracy (ACC). One specific area of improvement that we hope to address in the future is class imbalance. Although the F1-score is promising, we routinely saw a much higher precision than recall, this corresponds to the much larger amount of negative class instances (non-cancerous) compared to the smaller amount of positive class instances (cancerous). We also hope to spend more time finding the *sweet spot* of neural style transfer to achieve the best results on our augmented training set.

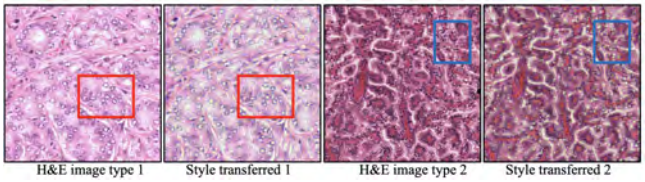

Figure 2: Examples of Neural Style Image Transform using Van Goh's *The Starry Night*

Examples of two types of neural-style transferred images are shown in Figure 2. The first H&E image on the left was transferred to a neural-style image creating some coarse in the overall region while the second H&E image on the right was transferred to the same style image but blurring each nuclei dense. Our experiment was performed by using all neural-style transferred images regardless of their texture representations.

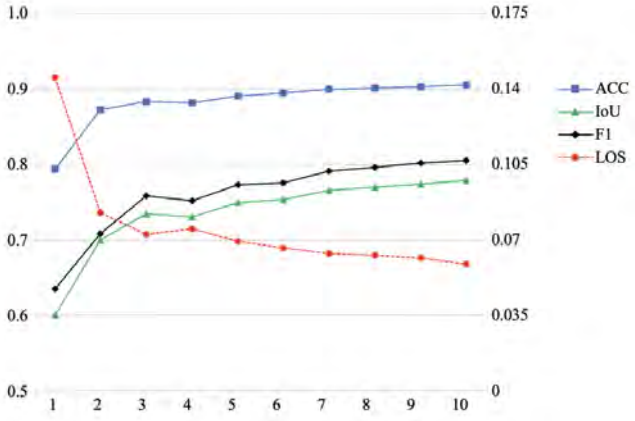

Figure 3: Performance Results on MonuSeg Datasest

The performance results on the MonuSeg dataset during training are shown in Figure 3. The x-axis represents the number of epochs used in the training. The left y-axis represents the range of ACC, IoU, and F1 scores and the right y-axis represents the range of loss (LOS) score. As the number of epochs increases, ACC represented as a box in blue converges quickly to around 0.9 while both IoU and F1 represented as a triangle in green and a diamond in black converge slightly slowly compared with the ACC to around 0.8. This is because both IoU and F1 begin converging at the lower scores, 0.6000 and 0.6348 respectively, while ACC begins converging at 0.7937. LOS continuously decreases but significantly between 1 and 3 epochs. The highest scores obtained by ACC, IoU, and F1

Table 1: Performance Results on TNBC Dataset

Models	F1
UNet++ EfficientNet-B0	0.5986
UNet++ EfficientNet-B1	**0.6785**
UNet++ EfficientNet-B2	0.5877
UNet++ EfficientNet-B3	0.6024
UNet++ EfficientNet-B4	0.6183
UNet++ EfficientNet-B5	0.5612
UNet VGG16 (Baseline)	0.5042
Proposed method	**0.6501**

are 0.9065, 0.7791, and 0.8058 respectively, while the lowest score obtained by LOS is 0.0587. Based on the experiment results in Figure 3, we can conclude that the combination of VGG16 and UNet, along with data augmentation techniques, generalizes effectively learning the proposed model through training on MonuSeg-2018 promising predictions on breast cancer images from the TNBC dataset.

The performance results on the TNBC dataset containing 50 images are shown in Table 1. We completed the training of the proposed model using the MonuSeg dataset and tested it on the TNBC dataset. The baseline model we defined for the basis of this experiment was UNet VGG16. We also compared the experiment results to Ding et al. and Lagree et al. As shown in Table 1, the proposed method outperformed the baseline and all but one, UNet++ EfficientNet-B1 in terms of F1 score. Although the F1 score of the proposed method (0.6501) is slightly less than UNet++ EfficientNet-B1 (0.6785), there is no doubt that our method significantly outperforms the baseline model (0.5042). Looking at the metrics on each prediction, there are several outliers on the low end for F1 specifically, which we believe would be solved by better addressing the class imbalance and optimizing the augmentation techniques.

The proposed method was applied to H&E images with various nuclei density levels. The examples of prediction results on different nuclei density levels are shown in Figure 4. A high-density nuclei image (top left) was predicted by the proposed method (top right) and was compared with a human-annotated image (top middle). Likewise, a middle-density nuclei image (middle left) and a low-density nuclei image (bottom left) were predicted by the proposed method and were compared with their human-annotated images. Regardless of nuclei-density level, the overall experiment results show that the proposed method is well-performed on the nuclei predictions by comparing with the human-annotated regions. However, nuclei regions with inconsistent boundaries produce inaccurate results of regions. For example, the predicted nuclei regions in the red box of the high-density nuclei image show many imprecise results, while the predicted nuclei regions in the red box of the middle or low-density nuclei image show seamless results compared with the one in the high-density nuclei image.

3.2 Additional Experiments

In addition to the experiments on the MonuSeg dataset and TNBC dataset using the proposed method, we perform two additional experiments with the modification of the original UNet. While the proposed method described in Section 3 uses both the VGG16

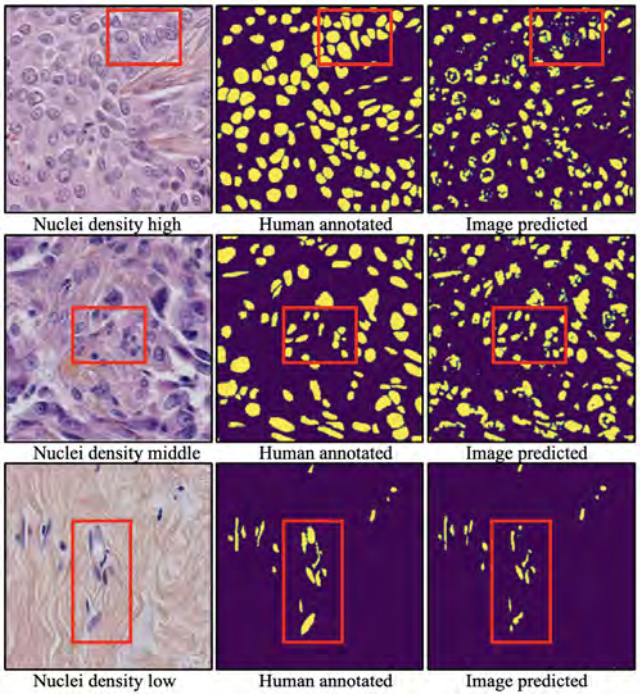

Figure 4: Examples of Prediction Results on Nuclei Density Levels

network and style-tranUNetsferred images, the modified UNet follows the original UNet architecture with different input sizes only because the main purpose of these experiments is to investigate whether the UNet can perform well on different types of cells such as blood, blood vessels, fat, glandular secretions, necrosis or debris, and plasma cells, as well as on different sizes of cancer-related cells such as a tumor, tumor-infiltrating lymphocytes, and stroma, or not, remaining the comparison between the proposed method and the UNet over the BCSS dataset in the future work.

3.2.1 Method. UNet is a convolutional neural network originally designed for image segmentation and the basic structure of the UNet architecture consists of two paths. The first path is about a contracting path which is similar to a regular convolution network and the second path is about an expansion path consisting of upsampling and concatenations with features from the contracting path. This expansion allows the network to learn localized classification information. Six different types of images of tissue cells: blood, glandular secretions, blood vessels, necrosis, fat, and plasma cells, were used for comprehensive assessment and fed to the UNet as the input sources. These images were extracted in 20x magnification from whole slide images (WSIs). WSIs are large-sized digital images, typically 1 or 2 GB per slide, created by digital equipment scanning glass slides used for examining tissue samples. These digitized images have been used for capturing, managing, and interpreting pathology information assisting traditional diagnostic processes. We also used three cancerous types of tissue cells, such as a tumor, stroma, and tumor-infiltrating lymphocytes (TILs), for

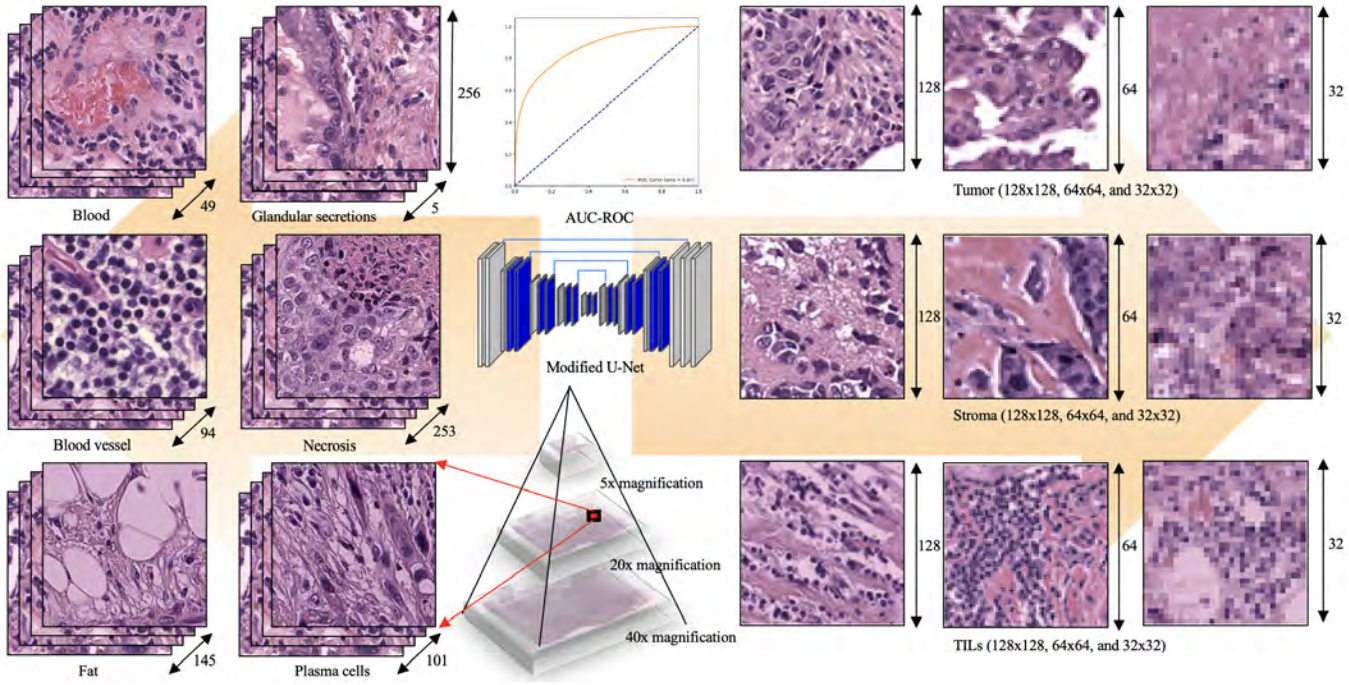

Figure 5: The Overview of Cancer-related Image Segmentation using UNet

the performance evaluation of the modified UNet with different input sizes.

The detailed overall process of the proposed method is shown in Figure 5. Digitized histopathology images are obtained at 20x magnification. 546 heterogeneous images with six tissue types extracted from the histopathology images are used for comprehensive access to breast cancer. 300 cancerous regions with three tissue types extracted from the histopathology images are used for the identification of cancer cells with different sizes.

3.2.2 Experiment Results. The dataset used in this experiment was obtained from the Breast Cancer Semantic Segmentation (BCSS) dataset [2]. BCSS dataset is a large-scale dataset that contains over 20,000 segmentation annotations of tissue regions consisting of 151 H&E breast cancer slides. The dataset has been annotated by an annotation review process involving senior pathologists, junior pathologists, and medical students. We extracted 546 heterogeneous regions of 256x256 size randomly selected from the 151 slides. We also extracted 300 cancerous cell images with different sizes of 128x128, 64x64, and 32x32.

The main purpose of the experiment in this section is to verify the effectiveness of the UNet on not a specific type of cell image but the various types of cell images. We modified the UNet with different input sizes and used the model for performance evaluation. Two well-known evaluation metrics: Accuracy (ACC.) and Area under the Receiver Operating Characteristic curve (AUC) were used for evaluating the performance of the UNet on heterogeneous types of cell predictions.

The performance results of six types of cell predictions using UNet are described in Table 2. The experiment results indicate that a modified UNet shows a high accuracy (0.9062) on fat identification, but it shows a low accuracy (0.6813) on necrosis identification. The performance results of three types of cell predictions with

Table 2: Performance Results on Six Types of Cells

Types of Cells	ACC	AUC
Blood	0.9412	0.6696
Blood vessels	**0.9421**	0.5276
Fat	0.9062	**0.9625**
Glandular secretions	0.8109	0.5981
Necrosis or Debris	0.6813	0.7266
Plasma cells	0.8796	0.7208

Table 3: Performance Results on Tumor, TILs, and Stroma with Different Sizes

Types	Size	5 Eps.	10 Eps.	20 Eps.	50 Eps.
Tumor	32x32	0.8935	0.9245	0.9418	**0.9918**
	64x64	0.9071	0.9084	0.9212	0.9753
	128x128	0.8856	0.9156	0.9301	0.9604
Stroma	32x32	0.7993	0.8215	0.8481	**0.9279**
	64x64	0.8184	0.8254	0.8499	0.9044
	128x128	0.8111	0.8330	0.8508	0.8829
TILs	32x32	0.9081	0.9393	0.9419	**0.9886**
	64x64	0.9218	0.9425	0.9238	0.9752
	128x128	0.9363	0.9294	0.9373	0.9550

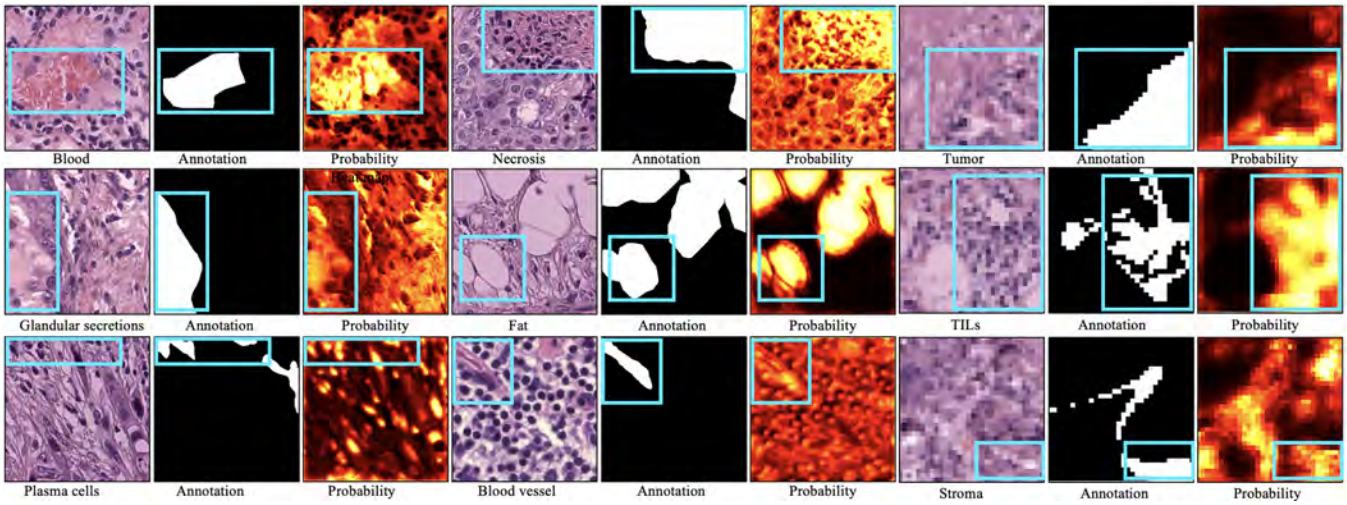

Figure 6: Results of the Comprehensive Experiments on Breast Cancer

different sizes using UNet are described in Table 3. The experiment results indicate that the smaller input size (32x32) of the tumor (0.9918), stroma (0.9279), and tumor-infiltrating lymphocyte (0.9886) shows better performance than the larger input size (128x128). The graphical view of the experiments is shown in Figure 6. The hematoxylin and eosin images of blood, the human annotation, and the probability map predicted by the modified UNet (top left). The hematoxylin and eosin image of necrosis, the human annotation, and the probability map (top middle). The hematoxylin and eosin image of the tumor with 32x32 size, the human annotation, and the probability map (top right inside). Glandular secretions, plasma cells, fat, blood vessels, TILs, and stroma follow the order of the above (middle and bottom). The probability heatmaps show that the identification of fat cells is intuitively very similar to the human annotation.

4 CONCLUSION

In this paper, we presented a new deep-learning method that combined VGG16 with UNet adopting neural-style networks for data augmentation to enhance the effectiveness of the model in histopathological image segmentation. A neural style transfer was performed to the H&E images along with various image augmentation techniques and was used as the input sources for the VGG16 encoder linked to the UNet decoder. The experiment results show that the proposed method reaches 90.00% accuracy, 77.91% intersection over union, and 80.58% F1 score on the MonuSeg dataset when training. Testing was done using the TNBC dataset and it shows that the proposed method reaches 65.01% F1 score which significantly outperforms the baseline model 50.42% but is less performed than UNet++ EfficientNet-B1 67.85%. In this regard, we plan to continue the experiment on another dataset, PathMNIST, as well as to use different deep learning models to investigate the effectiveness of the proposed method. Moreover, we presented various tissue-cell identification at the pixel level using a modified UNet. The experiment results show that the UNet is very effective on specific types of cells such as blood vessels (94.21% accuracy) and fat (96.25% area under the curve) but has a challenge on other types of cell predictions. In addition to the experiments on the various cell types, we performed additional experiments based on their sizes focusing on three cancer-related cells such as tumor, stroma, and TILs. The experiment results show that the smaller input size of the tumor (99.18%), stroma (92.79%), and tumor-infiltrating lymphocyte (98.86%) provides better performance than the larger input size.

ACKNOWLEDGMENTS

This research was partially supported by the National Science Foundation under Grant No. 2409704 and Grant No. 2409705, and Undergraduate Research and Creative Activities Grants at Kennesaw State University.

REFERENCES

[1] Ayat Abedalla, Malak Abdullah, Mahmoud Al-Ayyoub, and Elhadj Benkhelifa. 2020. The 2ST-UNet for Pneumothorax Segmentation in Chest X-Rays using ResNet34 as a Backbone for U-Net. *arXiv preprint arXiv:2009.02805* (2020).

[2] Mohamed Amgad, Habiba Elfandy, Hagar Hussein, Lamees A Atteya, Mai AT Elsebaie, Lamia S Abo Elnasr, Rokia A Sakr, Hazem SE Salem, Ahmed F Ismail, Anas M Saad, et al. 2019. Structured Crowdsourcing Enables Convolutional Segmentation of Histology Images. *Bioinformatics* 35, 18 (2019), 3461–3467.

[3] Vickie E Baracos, Tony Reiman, Marina Mourtzakis, Ioannis Gioulbasanis, and Sami Antoun. 2010. Body Composition in Patients with Non-Small Cell Lung Cancer: a Contemporary View of Cancer Cachexia with the Use of Computed Tomography Image Analysis. *The American journal of clinical nutrition* 91, 4 (2010), 1133S–1137S.

[4] Bhupender S Chhikara and Keykavous Parang. 2023. Global Cancer Statistics 2022: the Trends Projection Analysis. *Chemical Biology Letters* 10, 1 (2023), 451–451.

[5] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta, and Anil A Bharath. 2018. Generative Adversarial Networks: An Overview. *IEEE signal processing magazine* 35, 1 (2018), 53–65.

[6] Angel Cruz-Roa, Hannah Gilmore, Ajay Basavanhally, Michael Feldman, Shridar Ganesan, Natalie NC Shih, John Tomaszewski, Fabio A González, and Anant Madabhushi. 2017. Accurate and Reproducible Invasive Breast Cancer Detection in Whole-Slide Images: A Deep Learning Approach for Quantifying Tumor Extent. *Scientific reports* 7, 1 (2017), 46450.

[7] Yuxin Cui, Guiying Zhang, Zhonghao Liu, Zheng Xiong, and Jianjun Hu. 2019. A Deep Learning Algorithm for One-Step Contour Aware Nuclei Segmentation of Histopathology Images. *Medical & biological engineering & computing* 57 (2019), 2027–2043.

[8] Sergio Damas, Oscar Cordón, and Jose Santamaría. 2011. Medical Image Registration using Evolutionary Computation: An Experimental Survey. *IEEE Computational Intelligence Magazine* 6, 4 (2011), 26–42.

[9] Andre Esteva, Katherine Chou, Serena Yeung, Nikhil Naik, Ali Madani, Ali Mottaghi, Yun Liu, Eric Topol, Jeff Dean, and Richard Socher. 2021. Deep Learning-Enabled Medical Computer Vision. *NPJ digital medicine* 4, 1 (2021), 5.

[10] Arthur Francisco Araújo Fernandes, João Ricardo Rebouças Dórea, and Guilherme Jordão de Magalhães Rosa. 2020. Image Analysis and Computer Vision Applications in Animal Sciences: an Overview. *Frontiers in Veterinary Science* 7 (2020), 551269.

[11] Sourodip Ghosh, Aunkit Chaki, and KC Santosh. 2021. Improved U-Net Architecture with VGG-16 for Brain Tumor Segmentation. *Physical and Engineering Sciences in Medicine* 44, 3 (2021), 703–712.

[12] Intisar Rizwan I Haque and Jeremiah Neubert. 2020. Deep Learning Approaches to Biomedical Image Segmentation. *Informatics in Medicine Unlocked* 18 (2020), 100297.

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. Las Vegas, NV, USA, 770–778.

[14] Andreas Heinrich, Felix V Güttler, Sebastian Schenkl, Rebecca Wagner, and Ulf K-M Teichgräber. 2020. Automatic Human Identification based on Dental X-ray Radiographs using Computer Vision. *Scientific reports* 10, 1 (2020), 3801.

[15] Kirsi Holli, Anna-Leena Lääperi, Lara Harrison, Tiina Luukkaala, Terttu Toivonen, Pertti Ryymin, Prasun Dastidar, Seppo Soimakallio, and Hannu Eskola. 2010. Characterization of Breast Cancer Types by Texture Analysis of Magnetic Resonance Images. *Academic radiology* 17, 2 (2010), 135–141.

[16] Minui Hong, Jinwoo Choi, and Gunhee Kim. 2021. Stylemix: Separating Content and Style for Enhanced Data Augmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. Nashville, TN, USA, 14862–14870.

[17] Le Hou, Vu Nguyen, Ariel B Kanevsky, Dimitris Samaras, Tahsin M Kurc, Tianhao Zhao, Rajarsi R Gupta, Yi Gao, Wenjin Chen, David Foran, et al. 2019. Sparse Autoencoder for Unsupervised Nucleus Detection and Representation in Histopathology Images. *Pattern recognition* 86 (2019), 188–200.

[18] Aiyue Huang, Li Jiang, Jiangshan Zhang, and Qing Wang. 2022. Attention-VGG16-UNet: a Novel Deep Learning Approach for Automatic Segmentation of the Median Nerve in Ultrasound Images. *Quantitative Imaging in Medicine and Surgery* 12, 6 (2022), 3138.

[19] Philip TG Jackson, Amir Atapour Abarghouei, Stephen Bonner, Toby P Breckon, and Boguslaw Obara. 2019. Style Augmentation: Data Augmentation via Style Randomization. In *CVPR workshops*, Vol. 6. Long Beach, CA, USA, 10–11.

[20] Varun Jampani, Ujjwal, Jayanthi Sivaswamy, and Vivek Vaidya. 2012. Assessment of Computational Visual Attention Models on Medical Images. In *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing*. Mumbai, India, 1–8.

[21] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. 2017. A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology. *IEEE transactions on medical imaging* 36, 7 (2017), 1550–1560.

[22] Sanghoon Lee, Mohamed Amgad, Pooya Mobadersany, Matt McCormick, Brian P Pollack, Habiba Elfandy, Hagar Hussein, David A Gutman, and Lee AD Cooper. 2021. Interactive Classification of Whole-Slide Imaging Data for Cancer Researchers. *Cancer research* 81, 4 (2021), 1171–1177.

[23] Sanghoon Lee, Sunho Park, Jae-Ho Cheong, Sam C Wang, Matthew R Porembka, In-Ho Kim, Sung Hak Lee, and Tae Hyun Hwang. 2023. Spatial Image Analysis of Tumor-Infiltrating Lymphocytes on Gastric Cancer to Predict Anti-PD1 Inhibitor Response. *Cancer Research* 83, 7_Supplement (2023), 1360–1360.

[24] Shuying Liu and Weihong Deng. 2015. Very Deep Convolutional Neural Network based Image Classification using Small Training Sample Size. In *2015 3rd IAPR Asian conference on pattern recognition (ACPR)*. IEEE, Kuala Lumpur, Malaysia, 730–734.

[25] Anant Madabhushi and George Lee. 2016. Image Analysis and Machine Learning in Digital Pathology: Challenges and Opportunities. *Medical image analysis* 33 (2016), 170–175.

[26] Domingo Mery, Erick Svec, Marco Arias, Vladimir Riffo, Jose M Saavedra, and Sandipan Banerjee. 2016. Modern Computer Vision Techniques for X-ray Testing in Baggage Inspection. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47, 4 (2016), 682–692.

[27] Michael Nalisnik, Mohamed Amgad, Sanghoon Lee, Sameer H Halani, Jose Enrique Velazquez Vega, Daniel J Brat, David A Gutman, and Lee AD Cooper. 2017. Interactive Phenotyping of Large-scale Histology Imaging Data with HistomicsML. *Scientific reports* 7, 1 (2017), 14588.

[28] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. 2018. Segmentation of Nuclei in Histopathology Images by Deep Regression of the Distance Map. *IEEE transactions on medical imaging* 38, 2 (2018), 448–459.

[29] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. 2018. Attention U-net: Learning where to Look for the Pancreas. *arXiv preprint arXiv:1804.03999* (2018).

[30] Anindya Apriliyanti Pravitasari, Nur Iriawan, Mawanda Almuhayar, Taufik Azmi, Irhamah Irhamah, Kartika Fithriasari, Santi Wulan Purnami, and Widiana Ferriastuti. 2020. UNet-VGG16 with Transfer Learning for MRI-based Brain Tumor Segmentation. *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 18, 3 (2020), 1310–1318.

[31] Anindya Apriliyanti Pravitasari, Nur Iriawan, Ulfa Siti Nuraini, and Dwilaksana Abdullah Rasyid. 2022. On Comparing Optimizer of UNet-VGG16 Architecture for Brain Tumor Image Segmentation. In *Brain Tumor MRI Image Segmentation Using Deep Learning Techniques*, Jyotismita Chaki (Ed.). Academic Press, 197–215.

[32] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, Munich, Germany, 234–241.

[33] Devvi Sarwinda, Radifa Hilya Paradisa, Alhadi Bustamam, and Pinkie Anggia. 2021. Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer. *Procedia Computer Science* 179 (2021), 423–431.

[34] Dinggang Shen, Guorong Wu, and Heung-Il Suk. 2017. Deep Learning in Medical Image Analysis. *Annual review of biomedical engineering* 19 (2017), 221–248.

[35] Nahian Siddique, Sidike Paheding, Colin P Elkin, and Vijay Devabhaktuni. 2021. U-net and its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access* 9 (2021), 82031–82057.

[36] Khaoula Belhaj Soulami, Naima Kaabouch, Mohamed Nabil Saidi, and Ahmed Tamtaoui. 2021. Breast Cancer: One-Stage Automated Detection, Segmentation, and Classification of Digital Mammograms using UNet Model based-Semantic Segmentation. *Biomedical Signal Processing and Control* 66 (2021), 102481.

[37] Yi-Sheng Sun, Zhao Zhao, Zhang-Nv Yang, Fang Xu, Hang-Jing Lu, Zhi-Yong Zhu, Wen Shi, Jianmin Jiang, Ping-Ping Yao, and Han-Ping Zhu. 2017. Risk Factors and Preventions of Breast Cancer. *International journal of biological sciences* 13, 11 (2017), 1387.

[38] Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer Journal for Clinicians* 71, 3 (2021), 209–249.

[39] Bram Van Ginneken, Arnaud AA Setio, Colin Jacobs, and Francesco Ciompi. 2015. Off-the-shelf Convolutional Neural Network Features for Pulmonary Nodule Detection in Computed Tomography Scans. In *2015 IEEE 12th International symposium on biomedical imaging (ISBI)*. IEEE, Brooklyn, NY, USA, 286–289.

[40] Zhenqin Wu, Alexandro E Trevino, Eric Wu, Kyle Swanson, Honesty J Kim, H Blaize D'Angio, Ryan Preska, Gregory W Charville, Piero D Dalerba, Ann Marie Egloff, et al. 2022. Graph Deep Learning for the Characterization of Tumour Microenvironments from Spatial Protein Profiles in Tissue Specimens. *Nature Biomedical Engineering* 6, 12 (2022), 1435–1448.

[41] Yin Xu, Yan Li, and Byeong-Seok Shin. 2020. Medical Image Processing with Contextual Style Transfer. *Human-centric Computing and Information Sciences* 10, 1 (2020), 1–16.

[42] Qiao Zhang, Zhipeng Cui, Xiaoguang Niu, Shijie Geng, and Yu Qiao. 2017. Image Segmentation with Pyramid Dilated Convolution based on ResNet and U-Net. In *Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14-18, 2017, Proceedings, Part II 24*. Springer, Guangzhou, China, 364–372.