8-2024

Development of Feature Extraction Models to Improve Image Analysis Applications in Cancer

Yu Shi
University of Nebraska-Lincoln

Recommended Citation

Shi, Yu, "Development of Feature Extraction Models to Improve Image Analysis Applications in Cancer" (2024). *Dissertations and Doctoral Documents from University of Nebraska-Lincoln, 2023–*. 186.
https://digitalcommons.unl.edu/dissunl/186

DEVELOPMENT OF FEATURE EXTRACTION MODELS TO IMPROVE IMAGE ANALYSIS APPLICATIONS IN CANCER

by

Yu Shi

A DISSERTATION

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Doctor of Philosophy

Major: Complex Biosystems
(Systems Analysis)

Under the Supervision of Professor Chi Zhang

Lincoln, Nebraska
August, 2024

DEVELOPMENT OF FEATURE EXTRACTION MODELS TO IMPROVE IMAGE ANALYSIS APPLICATIONS IN CANCER

Yu Shi, Ph.D.
University of Nebraska, 2024

Adviser: Chi Zhang

Cancer poses a significant global health challenge. With an estimated 20 million new cases diagnosed worldwide in 2022 and 9.7 million fatalities attributable to the disease, the economic burden of cancer is immense. It impacts healthcare systems and imposes substantial costs for its care on patients and their families. Despite advancements in early detection, prevention, and treatment that have reduced overall cancer mortality rates, the growing prevalence of cancer, particularly among younger individuals, remains a pressing issue.

Recent advancements in medical imaging technology have progressed significantly with the help of emerging computer vision and artificial intelligence (AI) technology. Despite these advancements, medical imaging analysis in cancer research and clinical settings faces significant challenges. Analyzing data produced by sophisticated imaging technologies, such as CT or MRI, is still labor-intensive, limiting its usability and contributing to disparities in cancer care and data hungriness for researchers. AI-assisted analysis has the potential not only to reduce cost and turnover time but also to increase the accuracy of clinical applications. Furthermore, it provides opportunities to integrate various types of data and information for better prediction, benefiting both patients and physicians.

The research described in this dissertation aims to improve cancer imaging analysis by presenting the design and implementation of novel AI architectures. In this dissertation, I developed AI-based algorithms focused on two primary objectives. (1) **Develop feature extraction methods to improve model accuracy.** I applied advanced techniques to extract and learn critical image features associated with cancer prognosis to improve diagnostic tool accuracy and reliability. (2) **Develop advanced generative models to synthesize high-quality image data.** I developed deep-learning-based methods to learn latent representations and synthesize high-quality 3D images of tumor sites, facilitating better visualization and assessment of cancerous tissues.

This dissertation showcases the immense potential of AI in revolutionizing cancer diagnostics, providing a foundation for further research and development in this critical healthcare field. The proposed AI frameworks, incorporating innovative applications of machine learning and deep learning methods, will undoubtedly drive ongoing efforts to reduce cancer worldwide and tackle major challenges in this area.

DEDICATION

To My Beloved Parents: You gave me the roots to ground myself and the wings to soar. Your unwavering support, sacrifice, and belief in my dreams have been my anchor and compass. This achievement is not just my own but a testament to your enduring love and encouragement.

ACKNOWLEDGEMENTS

I extend my deepest gratitude to my advisor, Dr. Chi Zhang, whose expertise and insightful guidance have been instrumental in shaping my research and scholarly growth. I am immensely thankful for the encouragement and invaluable feedback from my committee members, Dr. Hasan Otu, Dr. Jennifer Clarke, and Dr. Etsuko Moriyama, whose perspectives have made this dissertation possible. Thanks to Dr. Maneesh Jain for adding clinical significance to my dissertation.

I am also grateful to my loved ones, who have provided me with unwavering support and companionship throughout this journey. Your belief in my abilities and constant encouragement has strengthened and motivated me.

Thank you all for your integral roles in this endeavor. I sincerely appreciate the wisdom, support, and friendship you have provided.

PREFACE

Parts of the contents discussed in this dissertation have been published as follows:

Chapter 2:

Shi, Yu, Ethan Wahle, Qian Du, Luke Krajewski, Xiaoying Liang, Sumin Zhou, Chi Zhang, Michael Baine, and Dandan Zheng. "Associations between statin/omega3 usage and MRI-based radiomics signatures in prostate cancer." *Diagnostics* 11, no. 1 (2021): 85. Used with permission.

Chapter 3:

Shi, Yu, Hannah Tang, Jianxin Sun, Xinyan Xie, Huijing Du, Dandan Zheng, Chi Zhang, and Hongfeng Yu. "Tissue-Specific Color Encoding and GAN Synthesis for Enhanced Medical Image Generation." In *2023 IEEE International Conference on Big Data (BigData)*, pp. 4344-4349. IEEE, 2023. Used with permission.

Chapter 4:

Shi, Yu, Hannah Tang, Michael J. Baine, Michael A. Hollingsworth, Huijing Du, Dandan Zheng, Chi Zhang, and Hongfeng Yu. "3DGAUnet: 3D generative adversarial networks with a 3D U-net based generator to achieve the accurate and effective synthesis of clinical tumor image data for pancreatic cancer." *Cancers* 15, no. 23 (2023): 5496. Used with permission.

Table of Contents

List of Tables

List of Figures

CHAPTER 1. INTRODUCTION

1. Motivation

Cancer significantly impacts human society both globally and within specific populations. In 2022, there were approximately 20 million new cancer cases diagnosed worldwide, leading to 9.7 million deaths [1]. The estimate is that new cases will reach 35 million by 2050, with the increasing number and aging of people on the planet. Cancer places substantial economic burdens on healthcare systems and societies. The burden of care, treatment, and palliation among cancer patients is significant both within public health resources and among individual cancer patients and their families. Advances in early detection, prevention, and treatment have led to a decline in overall cancer mortality rates; for instance, reducing smoking and improving cancer screening have saved millions of lives over the past few decades. But a growing prevalence of cancer, especially among younger people, is a constant challenge. So, it is evident that preventive interventions and equitable service delivery should be increased. These points make the complexity and multi-faceted impact of the disease on society clear while pointing out that despite many advancements, the fight against the disease is still very much a work in progress. Especially, CT and MRI imaging plays a very important role in cancer prevention, diagnosis, and treatment; there is a need to further improve medical image analysis and modeling for cancer research and clinical practice.

The current era is particularly exciting for biomedical research due to the convergence of big data generated by the advancement of experimental techniques. While vast amounts of complex data, including genomic, proteomic, and clinical datasets, have become available, there is an urgent need to establish sophisticated analytical methods to extract insights. Recent advancements in medical imaging technology have progressed significantly with the help of emerging computer vision and artificial intelligence (AI) technology. Despite these advancements, data analysis in cancer research, and clinical settings still faces significant challenges. Therefore, it is important to develop AI models and data analytical methods. This synergy allows for the development of more research tools and a better understanding of cancer, which is this thesis's core objective.

This chapter summarizes the background and current status of medical imaging for cancer and computer vision applications in medical imaging analysis. An overview of the AI technologies applied in medical imaging analysis, especially related to cancer research, diagnosis, and treatment, is also provided. There are gaps between the advances in AI method development and their applications in cancer imaging. This dissertation focuses on three major gaps in existing research. Firstly, the associations between quantitative image features and clinical events are largely unknown (e.g., patient outcome and prognosis). Secondly, the methods for synthesizing high-quality medical imaging data, particularly for 3D volumetric data, are not yet fully mature, highlighting the current limitations in the field. Last but not least, there is room for improvement in model accuracy by integrating multi-modal data. *To solve these problems, the research described in this dissertation aims to develop a novel image feature extraction method to connect clinical events and improve model accuracy, as well as to develop an advanced generative model to synthesize high-quality image data.*

2. Medical Imaging for Cancer

2.1. Medical imaging in cancer research, diagnosis, and treatment

Medical imaging has been a cornerstone in cancer research, diagnosis, and treatment. It is a fundamental tool in the diagnosis of cancer. It contains essential information for clinical processes to enhance decision-making results. Several imaging modalities are currently used to detect, characterize, and stage various cancers. Computed Tomography (CT), Positron Emission Tomography (PET), and Magnetic Resonance Imaging (MRI) have become more popular because they provide the highest sensitivity or most detailed specificity. Medical imaging is vital for diagnosing diseases as well as treatment planning, targeting, and surveillance. Over the past decade, with the help of emerging computer vision and AI technology advancements, technologies for medical imaging have progressed significantly, contributing to improved detection, diagnosis, and treatment of cancer.

The use of medical imaging is highly recommended in almost all steps related to the care of a person with cancer, from early detection to treatment planning and follow-up. It is critical because it directly connects with the patient's outcomes and the effectiveness of cancer therapies. With improvements in image quality and reduction in radiation exposure, combined with the development of improved AI technologies, better and even more personalized cancer care can be achieved.

2.2. Existing challenges in medical imaging for cancer.

Even though medical imaging has improved vastly, it still faces challenges. For example, sophisticated imaging technologies such as PET/MRI are also expensive to establish in many health facilities, limiting access to most patients. The costs of such equipment and analysis services are disproportionately high in healthcare systems, contributing to disparities in cancer care outcomes [2].

With the increase in data generated from various imaging modalities—e.g., MRI, PET, and CT—integrating the data from multi-modal sources is essential for a holistic understanding of tumor biology and proper diagnosis and treatment planning [3]. However, it requires a complex process, and currently, some gaps need to be filled. While medical imaging is one area where AI is making giant steps in improving medical outcomes, it has been raising deep concerns about its ethics and operation. Significant problems are associated with data privacy and potential bias in AI algorithms. Clear standards and regulation of AI are needed for its safe and effective use in medical diagnostics [4].

Furthermore, radiologists from different institutions interpret imaging results differently, resulting in conflicting diagnoses and treatment plans. Standardizing interpretation protocols is necessary to reduce such variability. With the adaptation of AI technologies, especially for deep models, model explainability is also important in risk and quality control [5].

3. Computer Vision Application in Medical Imaging Analysis

Computer vision is a field of AI that allows machines to interpret and make decisions based on visual data. It involves creating algorithms and models to process, analyze, and comprehend images and videos, replicating human vision [6]. This section provides a broad overview of this field of research and sets the technical foundation of the following chapters.

3.1. Image clustering

Image clustering in medical imaging is a critical area of research that enhances diagnostic processes and supports clinical decision-making. Clustering techniques group similar images or image regions, aiding in identifying patterns and anomalies essential for diagnosing various medical conditions. Here are some key aspects of image clustering in medical imaging:

Feature Extraction: Effective clustering starts with robust feature extraction. Typically, medical images are represented in a lower-dimensional space that is conducive to clustering using techniques such as texture analysis, edge detection, and shape-based features.

Dimensionality Reduction: Given the high dimensionality of medical images, techniques such as Principal Component Analysis (PCA) [7] and autoencoders [8] are employed to reduce the dimensionality of feature spaces before clustering.

Clustering Algorithms: The specific requirements and nature of the medical imaging data dictate the use of a variety of clustering algorithms. Hierarchical clustering and K-means clustering are prevalent algorithms.

Clustering in medical imaging has been applied in numerous areas, such as tumor detection, brain imaging [9], and tumor segmentation [10].

3.2. Image classification

Image classification is a supervised learning technique in computer vision. A classification system learns to recognize and categorize images based on specific classes. This is the process of analyzing and tagging features within the image. Classifications are done in predefined categories, where images are placed based on the similarity of their content. Image classification is also essential for automated diagnosis systems where algorithms for detecting and classifying medical conditions are applied for X-ray, MRI, and CT images [11]. Automating the task of disease classification using medical images can assist in screening large populations quickly, consistently, and accurately. Image-based disease classification enables the early detection of diseases, hence increasing the prognosis by early treatment. The main elements of image classification include:

Training Data: A large dataset with labeled images is required to train an image classification model. Each image in this dataset is labeled with a single or multiple categories. Using these training datasets, the model learns how to identify patterns characteristic of each category.

Positive Dataset: A positive dataset refers to data points representing the condition or class of interest. In medical imaging, this typically includes images or cases containing the pathology, feature, or condition that the model needs to learn to detect or diagnose. For example, in a study aimed at identifying tumors, the positive dataset would consist of images that contain tumors.

Negative Dataset: Conversely, a negative dataset consists of data points that do not display the condition or class the model is targeting. These are essentially control samples used to teach the model what does not constitute a positive case. In the tumor identification example, the negative dataset would comprise images without tumors.

Model Types: Model types commonly used for image classification mainly come from traditional machine learning approaches, such as support vector machine (SVM), decision trees, random forests (RF), and k-nearest neighbors (KNN). In recent years, deep learning has revolutionized the field. Key models that have driven deep learning technologies are convolutional neural networks (CNNs), transformers, and autoencoders, all of which can achieve automated feature extraction and perform well on complex datasets.

Feature Extraction: The typical first step of image classification is to extract relevant features from an image. The features might be related to shapes, textures, or intensity patterns characteristic of some medical conditions.

Model Training: The extracted features are used to train a model, most typically through machine learning or deep learning algorithms. CNNs are extensively used in deep learning because they can learn hierarchical representation in images, which fits very well for image classification.

Classification: After the training is done, the model classifies new images into various classes such as ‘normal’ or ‘abnormal’. It may further classify the images based on learned features into specific diseases such as pneumonia, multiple kinds of cancer, or other pathological conditions.

3.3. Object detection

Object detection can be seen as one application of computer vision technologies in medical field. It is used to identify and localize several structures or abnormalities in the given medical images, such as X-rays, MRI scans, CT scans, and ultrasound images. This technology made it possible to diagnose diseases, schedule treatment, and evaluate patients' progress much faster and more accurately. The key components are:

Preprocessing: Images often require enhancement for quality and visibility. Some treatments that may need to be implemented include noise removal, enhancement for contrast, and standardization of dimensions and scales among all images.

Object Detection Algorithms: These algorithms are dedicated to analyzing any presented medical images in detecting an object that carries some meaning for diagnosis and treatment. Standard algorithms include You Only Look Once (YOLO) [12], transformers, CNNs, and CNN variations.

Annotation and Localization: The detected objects are mostly annotated with bounding boxes indicating the coordinates of an object within an image. This aids localization and indicates an anomaly's exact location, size, and shape.

Validation and Interpretation: Once objects have been detected, medical practitioners will analyze them for accurate interpretation. This is an important step that combines machine efficiency with human expertise to arrive at diagnostic decisions.

3.4. Segmentation

Segmentation in medical imaging is the process that allows the division of images into segments to simplify or change the representation of an image into something more meaningful and more accessible for computational image analysis. It is also important to demarcate the regions of medical interests, such as organs or tumors, in medical images such as CT and MRI scans or ultrasounds. Segmentation plays a vital role in medical imaging by accurately representing and analyzing anatomical and pathological features. More precise segmentation would also provide better data quality for other CV tasks, such as image classification. Here is an overview of the process:

Preprocessing: Noise reduction, contrast enhancement, and normalization are common preprocessing steps to improve the quality of images before segmentation. The preprocessing typically depends on the image modality and the anatomical characteristics of the object.

Segmentation Techniques: Traditional segmentation techniques are signal processing methods such as thresholding, region growing, and watershed segmentation. In the past decade, machine learning and deep learning-based methods have started to be applied, with more advanced methods incorporating algorithms into practices such as U-Nets or CNNs to ensure accuracy and automation during the segmentation processes.

Postprocessing: Smoothing edges, eliminating small spurious objects, or filling holes in the segmented regions are standard postprocessing practices.

Analysis and Interpretation: Analysis of the images to reflect on medical conditions is done by medical professionals or automated quantitative analysis tools.

3.5. Image registration

Image registration transforms images into the same reference coordinate system in medical imaging. It aligns two or more photos of the same scene taken at different times, from various viewpoints, or by other sensors. Combining information from multiple images helps provide a more comprehensive analysis of the medical field [13]. Here is an overview of the process:

Preprocessing: Images can be preprocessed to enhance quality, improve features, and standardize formats so that they can be more effectively registered.

Feature Detection and Matching: Edge features, contours, or other anatomical landmarks are identified and used to match corresponding points between images.

Transformation Model Estimation: The mathematical model that should be selected can be either rigid, affine, or non-rigid to best suit the nature of the alignment. The parameters should be adapted to align features between the images optimally.

Resampling and Interpolation: Once aligned, the images are commonly resampled using interpolation techniques to build up the final image while preserving the actual pixel values from the original images.

Validation: Registration quality is assessed using quantitative metrics or by expert inspection based on visualization to ensure correct alignment.

3.6. Radiomic feature extraction

Radiomics, which extracts quantitative features from the images, significantly increases the throughput of the quantifiable representation of each image screening, which ensures better differentiation with statistical analysis. Recent advances in radiomics make it possible to extract features from images [14]. Feature extraction calculates quantifiable characteristics of the signal level on segmented images. These features can be broadly categorized into first-order statistics, shape-based features, texture features, and higher-order statistics (**Figure 1.1**)

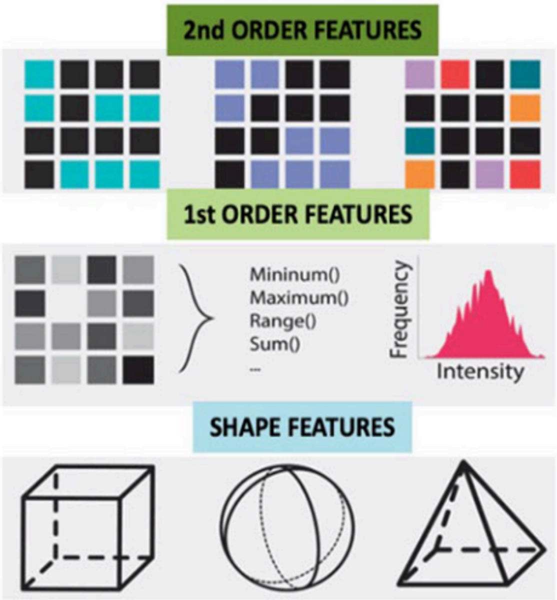


Figure 1.1 Schematic View of Radiomics [15].

An independent international collaboration called the Image Biomarker Standardization Initiative (IBSI) aims to standardize medical image feature extraction protocol and provide image biomarker nomenclature and definitions. IBSI guidelines are followed in the research in this dissertation. Studies that use quantifiable features extracted with radiomics techniques show promising capabilities in predicting cancer outcomes [16], [17]. Based on the guidelines of IBSI, hundreds of image features can be derived from one image [18]. For classification tasks, the number of features should be regulated according to the learning task's complexity. Since radiomic features could be derived with closely related operations, using all of the radiomic features directly may cause overfitting. Therefore, most radiomics studies require dedicated feature dimension reduction to find the balance between specificity and generality. This causes low repeatability and low reproducibility of radiomic features.

3.7. Challenges for application of computer vision in medical imaging tasks

In the rapidly evolving field of medical imaging, technological advancements bring exciting prospects for improved diagnostic accuracy and patient outcomes. However, effectively implementing these technologies presents challenges that must be addressed to harness their full potential. Here is a summarized overview of existing challenges in medical imaging in terms of computer vision application [19]:

*Data Quality and Quantity***:** The effectiveness of image classification heavily depends on having access to high-quality and extensively annotated datasets. Obtaining such data can be expensive and hence, such data are often scarce [20].

Generalization: Models must be generalized appropriately, avoiding overfitting the specific characteristics of the training dataset, so that they can be applied effectively to new, unseen data.

Data Variability: The diversity in medical imaging due to different modalities, settings, demographics, and disease manifestations often enhances the robustness of classification models.

Data Privacy and Security: Secure handling of patient data is essential to protect privacy when training algorithms [21].

Model Accuracy: In medical contexts, accuracy and reliability are vital due to the severe implications of false positives and negatives.

Integration into Clinical Workflows: Technologies must integrate seamlessly into existing medical workflows without necessitating significant changes or extensive staff retraining.

Computational Demand and Efficiency: Advanced techniques, particularly those involving deep learning, require substantial computational resources and expertise, and the process can be computationally intensive, especially with high-resolution images or complex models.

Standardization and Reproducibility: The variability in results produced by different algorithms poses challenges for standardization and reproducibility in clinical settings. This is compounded by variability in imaging equipment and protocols across various institutions [22].

Integration into Clinical Practice: There is a significant challenge in translating research findings into practical clinical applications, necessitating standardized protocols for image acquisition, feature extraction, and analysis to ensure findings are generalizable and applicable across different studies.

4. Association Study in Medical Imaging

Association studies in medical imaging are a crucial research approach used to identify correlations between imaging features and various clinical outcomes, including disease characteristics, prognosis, and response to treatment. These studies help understand how certain visual markers or patterns observed in medical images relate to specific health conditions or genetic traits. Radiomics to extract quantitative image features combined with machine learning models are commonly used to draw the association between specific image features and clinical variables or biological processes [23].

Association studies can help identify imaging biomarkers that predict disease severity, progression, and outcome. For example, certain features in lung CT images may be associated with faster disease progression in idiopathic pulmonary fibrosis [24]. Researchers may find associations between mutations in specific genes and radiographic features. Breast cancer MRI phenotypes could be related to their underlying molecular biology, revealed by using RNA sequencing, which can help understand the genetic drivers of disease. Moreover, the MRI features may be used as imaging biomarkers for drugs targeting the ribosome [25]. Image-based association studies, particularly in medical imaging and disease diagnosis, are a relatively young field within biomedical research, with numerous challenges that need to be addressed. For example, here are some of the current challenges:

High Dimensionality: Medical images can involve high dimensions of data, resulting in complexity in statistical analysis and often calling for advanced dimensionality reduction techniques.

Validation and Reproducibility: Ensuring validity in different populations and imaging settings is demanding yet essential for clinical implementation.

Integration with Clinical Data: Combining imaging data with clinical data requires robust data integration platforms, and there are compatibility issues regarding data privacy.

5. AI Techniques Used in Medical Imaging

AI techniques have become increasingly important in medical imaging analysis, offering transformative potential across various aspects of healthcare. Some essential AI techniques and their applications in medical imaging are described below:

5.1. Support Vector Machine (SVM)

SVM is a popular supervised learning algorithm for classification and regression tasks. It performs well in high-dimensional spaces and is particularly useful when the number of dimensions exceeds the number of samples.

SVM operates on the principle of finding the optimal hyperplane to separate two different classes with maximum margin. This hyperplane is the decision boundary that separates one class of data points from another. Critical components of SVM are as follows:

Hyperplane: In SVM, a hyperplane is a flat affine subspace (affine subspace, which is one dimension less than the feature space) that divides the data points of different classes. In two-dimensional space, this is a line.

Margin: The margin is the distance between the hyperplane and the nearest data points from each class. Maximizing this margin is crucial for improving the model's generalization ability.

Support Vectors: These are the data points closest to the hyperplane, influencing its position and orientation. Support vectors are vital as they define the margin.

Linear SVM: In its simplest form, when the data is linearly separable (i.e., two classes can be separated by a straight line in two dimensions or a hyperplane in higher dimensions), SVM finds the hyperplane that maximizes the distance between the classes. The optimization problem is typically solved with quadratic programming.

Kernel Trick: The kernel trick uses a linear classifier to solve a non-linear problem by mapping the input space into a higher-dimensional feature space, computed beforehand, called kernels. The kernel function calculates the dot product of all data pairs in this higher-dimensional space. This allows SVM to fit the maximal margin hyperplane in the new, higher-dimensional feature space. For example, the radial basis function (RBF) kernel adds new features that measure the similarity (or distance) between data points and specific landmarks in the input space, leading to a smoother decision boundary.

SVMs are widely used in medical imaging for tasks such as classifying tissue types, identifying disease states, and differentiating between benign and malignant tumors [26], [27]. Their ability to handle high-dimensional data makes them suitable for complex and multidimensional image data.

5.2. Ensemble learning

5.2.1. Random Forest (RF)

RF is an ensemble technique where decision trees are built during training that outputs the class; the final output is the mode of the classes or mean prediction of the individual trees. This method is prevalent in the case of medical imaging as it can work well with large and complex datasets with a high-dimensional space, which is the case found in medical images. The key components are as follows:

Decision Trees: RF uses a large number of decision trees to make predictions. Each tree is a series of questions about the data that lead to a decision, predicting the outcome based on the data features.

Bootstrap Aggregating (Bagging): RF uses the Bagging technique to improve model accuracy. This involves creating numerous bootstrap samples of the original dataset (resampling) to train each tree in the forest independently. This process helps reduce variance and avoid overfitting.

Training Trees: From the training data, each decision tree in the forest is trained independently on a different bootstrap sample. The variety among the trees is achieved by using various samples and randomly selecting a subset of specified features for splitting at each node of the trees.

Random Feature Sampling: During each decision tree training split, RF does not look for the best feature among all features but looks for the best feature among a random subset of features. This results in many trees, reducing the risk of overfitting and making the model less likely to memorize any specific part of the data.

Ensemble of Predictions: Once all the trees are trained, RF makes a decision by combining the predictions of all the trees built. Classifications are usually done based on majority voting (the class commonly found among a larger number of trees is used). For regression, typically averaging of the predictions of all trees is used.

RF algorithms are crucial in classifying medical images into different diagnostic categories. For instance, RF can be utilized to classify breast cancer biopsy from mammograms, which is a critical factor in assessing the risk of breast cancer [28].

Additionally, RF is used for image segmentation, which involves dividing an image into segments representing different anatomical structures, such as brain tissues or tumors [29]. This segmentation is essential for more precise medical interventions.

5.2.2. Gradient Boosting Machine (GBM)

GBM is a sequential ensemble learning method for constructing a model step-wise, where each new model corrects errors made by existing models. The key idea behind GBM is to build new base learners, which predict the residuals or errors of prior models, and then add those models to minimize the overall prediction error. The boosting algorithm constructs this model stage-wise, which differs from how RFs build the model. GBM adjusts the weights of incorrectly predicted cases, making the model sensitive to the more complex cases, thereby improving the model's performance iteratively. This iterative improvement allows the model to focus more on the challenging cases, enhancing overall accuracy and robustness. By continuously refining the model's predictions, GBM achieves better performance over time, but when the problem is simple, it is prone to overfitting. GBM has been applied to various tasks in medical imaging; with its tree-based decision-making, it is a favorable method for risk stratification analysis, such as using imaging data to predict breast cancer patient recurrence risk [30]. GBM also can be used to segment brain tumors from MRI scans. The approach involved using GBM to segment the brain scans into cancer or healthy tissue and provide refined tumor labels, providing finer granularity [31].

5.3. Convolutional Neural Networks (CNNs)

CNNs are a class of deep neural networks widely used in analyzing visual imagery. CNNs are particularly powerful for tasks involving image recognition, object detection, and similar applications due to their architecture, which mimics how the human visual cortex works. From the input image to the output, CNNs are built into layers (**Figure 1.2**):

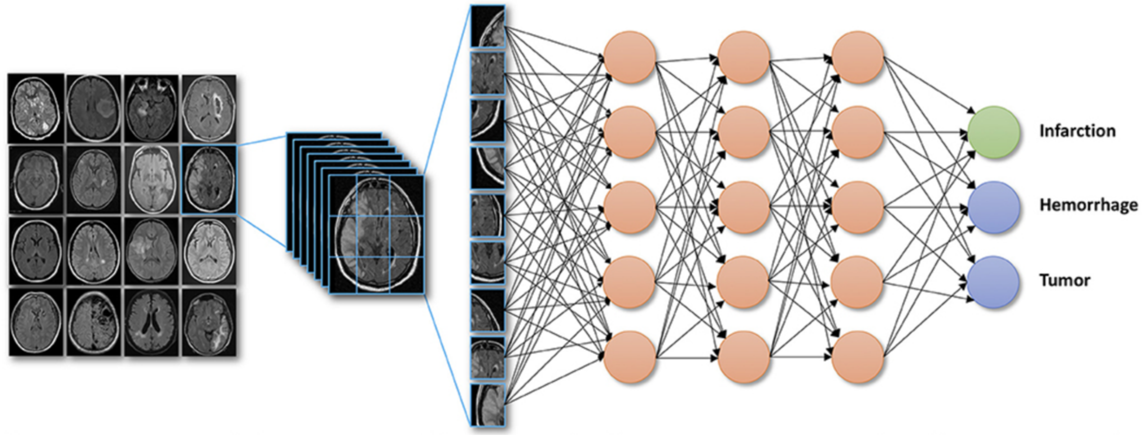


Figure 1.2 CNN Classifier [32].

Convolutional Layer: Convolutional layers employ multiple filters or kernels that perform matrix multiplication on the input image to extract features, transforming the image data from basic pixel-level information to higher-order features. These layers capture spatial hierarchies within images by utilizing several square matrices and generating feature maps.

Pooling Layer: Pooling layers are designed to decrease computational costs by reducing the dimensions of feature maps from convolutional layers. They perform pooling operations independently on each feature map. Common types of pooling include max pooling and average pooling. Pooling helps summarize the features detected in previous layers. Max pooling takes the maximum value from each cluster of neurons at the prior layer, effectively capturing the most prominent feature in a local patch of the feature map. Max pooling helps detect features that are invariant to small shifts and distortions. Contrary to max pooling, average pooling calculates the average values in each cluster of neurons. This method is less common as it tends to blur feature representations, diluting the presence of solid features but providing a more comprehensive summary of the features in the patch.

Fully connected layer: A fully connected layer, also known as a dense layer, is essential in a CNN for synthesizing the features extracted from previous layers—specifically, the convolutional and pooling layers. Positioned typically towards the end of the network, it consolidates these learned features into high-level reasoning about the input data, facilitating the network's decision-making or classification tasks. The number of neurons in the dense layer equals the number of classes in the classification case.

Activation Function: Activation functions in CNNs are vital to infuse non-linearity within the network, which is an absolute requirement for learning and performing complex tasks such as image recognition and classification. Detailed looks at some of the most commonly used activation functions in CNNs are as follows:

Rectified Linear Unit (ReLU):

$$f(x) = max(0, x) \tag{1}$$

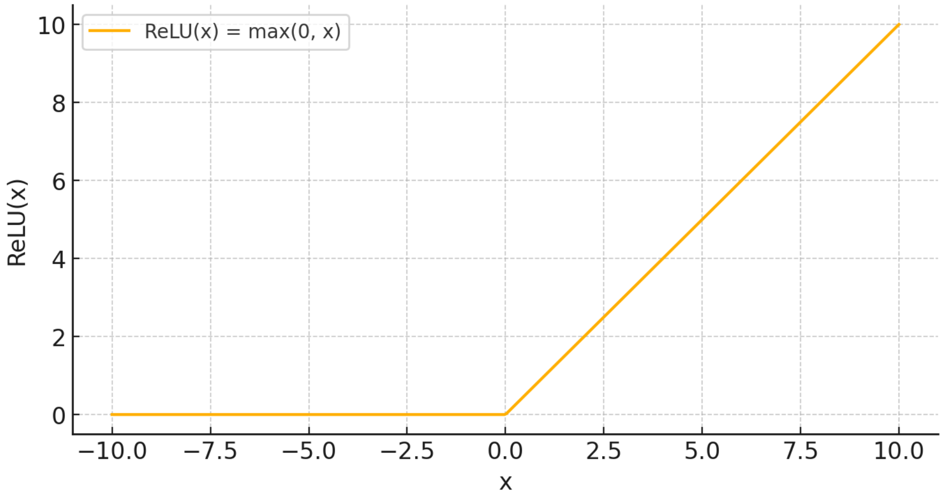


Figure 1.3 ReLU Activation Function.

As you can see in **Figure 1.3**, the function outputs 0 for all negative values and simply outputs the input variable x for all non-negative values. This simple yet effective behavior helps maintain the non-linear properties needed for learning complex patterns in data.

Leaky ReLU:

$$f(x) = \max(\alpha x, x) \tag{2}$$

where α is a small constant.

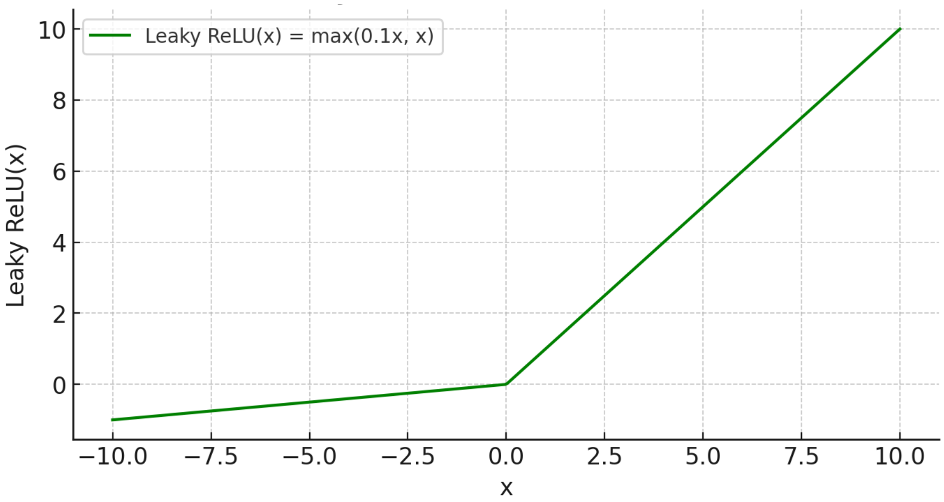


Figure 1.4 Leaky ReLU Activation Function.

It is similar to ReLU but has a slight positive slope for negative values (α=0.1 is used in **Figure 1.4**). This modification helps to keep some gradient flow during the training process. When the input values are less than zero, it prevents neurons from dying out, a common issue with the standard ReLU function.

Sigmoid:

$$f(x) = \frac{1}{1+e^{-x}} \quad (3)$$

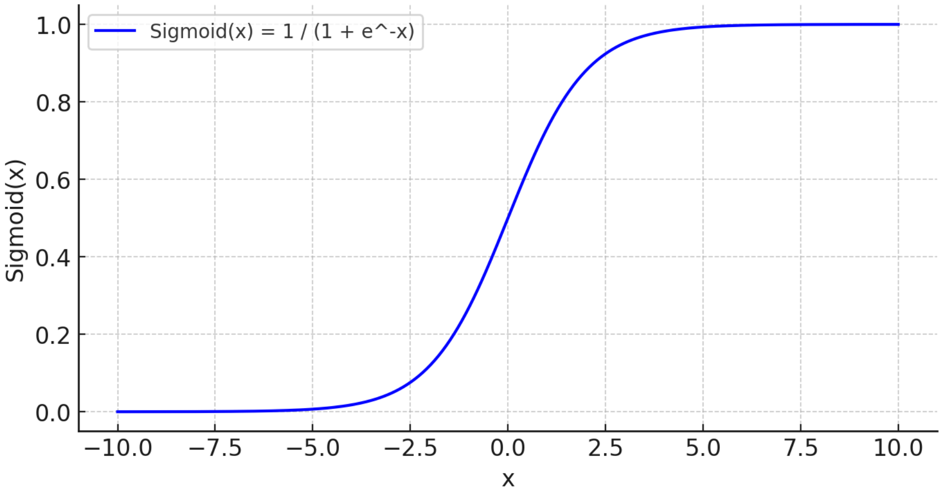


Figure 1.5 Sigmoid Activation Function.

The Sigmoid function outputs values between 0 and 1, making it especially useful for models that need to predict probabilities as outputs. As shown in **Figure 1.5**, The function smoothly increases from 0 to 1 and is centered at $x = 0$, where f(x) equals 0.5.

Tanh (Hyperbolic Tangent):

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{4}$$

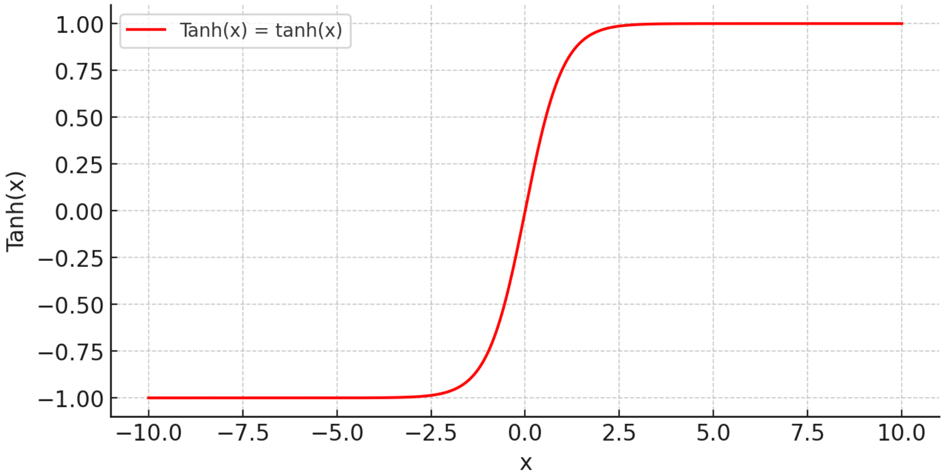


Figure 1.6 Tanh Activation Function.

The Tanh function outputs values between -1 and 1. As shown in **Figure 1.6**, it is an S-shaped curve similar to a sigmoid function but is zero-centered, which generally helps improve the convergence of the learning process during training. This function is especially beneficial when the model needs to differentiate between two symmetrically balanced classes around zero.

Softmax:

$$Softmax(z)_i = \frac{e^{z_i}}{\sum_{k=1}^{K} e^{z_k}} \tag{5}$$

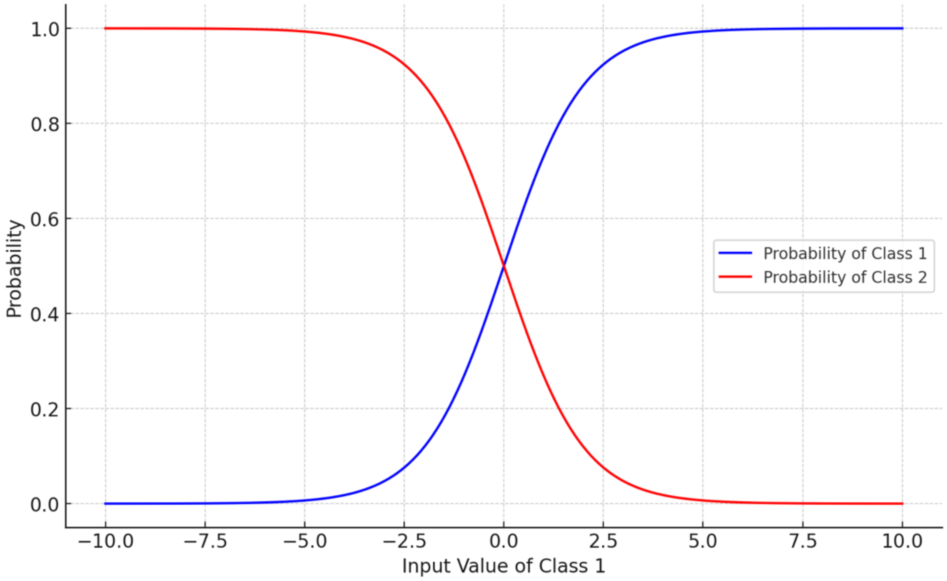


Figure 1.7 Softmax Activation Function.

In **Figure 1.7**, we see a simplified Softmax scenario with a binary classification task. As the input for Class 1 increases, the probability that the input belongs to Class 1 increases, illustrating how the Softmax function dynamically allocates probabilities between the two classes based on the input values. This visualization captures the essence of Softmax in a binary setting where the function ensures that the total probability for both classes sums to one. Softmax is primarily used in an output layer of a neural network to perform multi-class classifications, where it returns probabilities of each class, with the total summing to one.

5.4. U-Net

U-Net is one of the CNN architectures that has emerged in different image segmentation tasks, especially in medical imaging [33]. The network has a symmetric U-shaped architecture composed of a contracting path (encoder) and an expanding path (decoder) (**Figure 1.8**). The contracting path captures context and reduces the spatial dimensions of the image, while the expanding path enables precise localization and upscaling to the original image size. U-Net has become one of the most commonly used and effective architectures for performing semantic segmentation. The key features of U-Net that are different from CNNs are as follows:

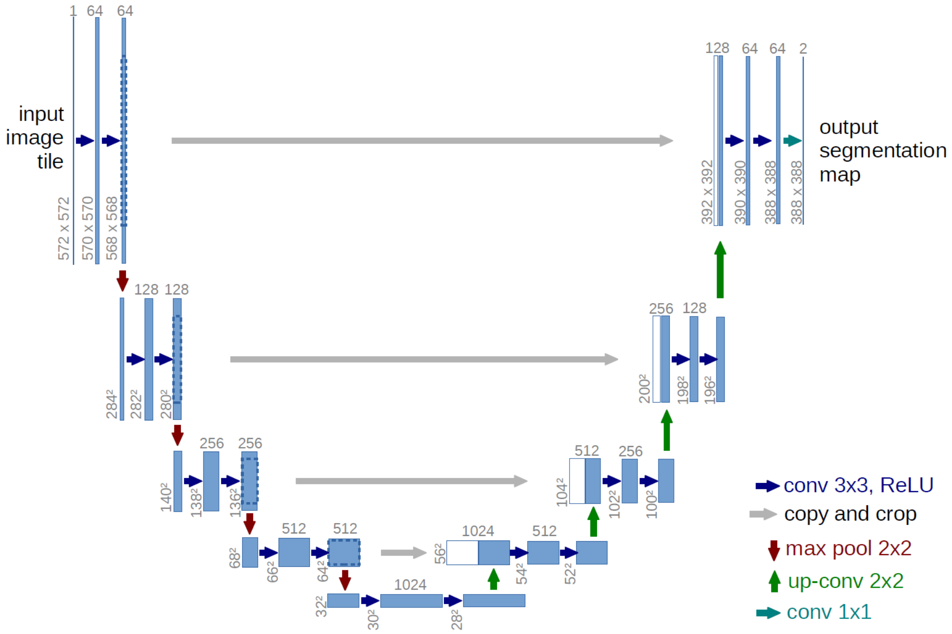


Figure 1.8 U-Net Architecture [33].

Contracting Path (Encoder): The contracting path consists of convolutional layers, each followed by a ReLU activation and a max-pooling operation for down-sampling. Each down-sampling step doubles the number of feature channels, allowing the network to learn more complex representations.

Bottleneck: The bottleneck is the deepest part of the network and connects the contracting and expanding paths. It consists of two convolutional layers followed by ReLU activations. This part of the network learns the most abstract features of the input image.

Expanding Path (Decoder): The expanding path consists of up-convolutional layers (transposed convolutions) that increase the spatial dimensions. Each up-convolution is followed by concatenation with the corresponding feature map from the contracting path (skip connections) and a set of convolutional layers with ReLU activations.

Skip Connections: Skip connections between the contracting and expanding paths, allowing the network to reuse features from the down-sampling path, avoiding gradient decay, and improving the precision of the segmentation. These connections enable the network to combine high-resolution spatial information with the contextual information learned during down-sampling.

Output Layer: The final layer is typically a 1×1 convolution that maps the feature representation to the desired number of output classes for segmentation.

5.5. Generative Adversarial Networks (GANs)

GANs [34] are a class of models comprising generator and discriminator models. They are trained simultaneously and optimally in a game-theoretic framework. The result is that the generator learns to produce data indistinguishable from accurate data, making

GANs exceedingly strong at generative tasks (**Figure 1.9**). Components of GANs are as follows:

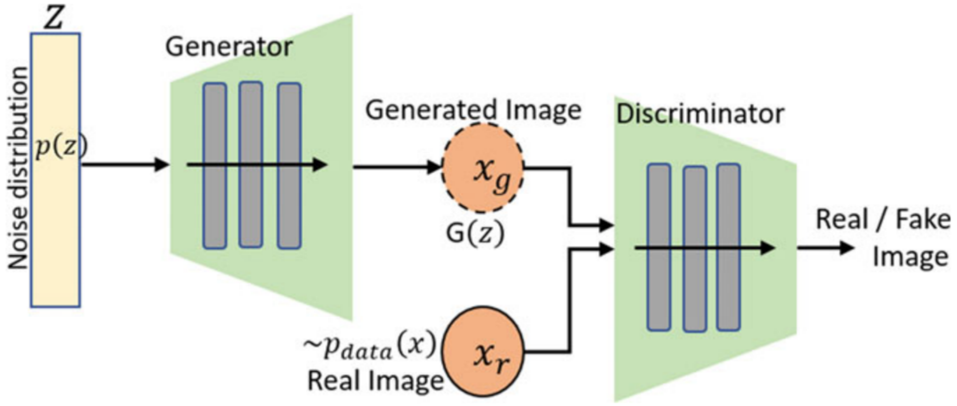


Figure 1.9 GAN Architecture [26].

Generator (G): The generator creates data that resembles the ground truth. It takes a random noise vector and transforms it into a data instance (e.g., an image). The generator aims to fool the discriminator into believing the generated data is accurate. *Discriminator (D):* The discriminator distinguishes between ground truth (from the training set) and synthetic data (produced by the generator). It outputs a probability value indicating whether the input data is real or fake.

Training Process: The generator and discriminator are trained simultaneously but with opposing objectives. The generator tries to minimize the probability of the discriminator correctly identifying fake data while the discriminator tries to maximize its accuracy.

Adversarial Training: The discriminator is trained on both real and fake data. It learns to assign a high probability to real data and a low likelihood to counterfeit data.

The generator is trained to produce data that maximizes the discriminator's output probability for fake data.

Optimization: The training process of GANs involves two loss functions: one for the discriminator loss and another for the generator loss. The discriminator loss tries to maximize the log probability of correctly classifying real and fake data, effectively learning to distinguish between authentic and generated samples. On the other hand, the generator loss is the log probability for which the discriminator correctly identifies the generated data as fake. Hence, it further pushes the generator to produce more realistic outputs. This is the adversarial process in which the generator and the discriminator are trained simultaneously but with opposing objectives. The generator keeps getting better at ways to make this hard—to produce data from some distribution that is indistinguishable from an actual data distribution. Mathematically, this can be expressed as:

$$\min_{G} \max_{D} V\,(D, G) = E_{x\sim p_{data}(x)}[\log D\,(x)] + E_{z\sim p_z(z)}\left[\log\left(1 - D\big(G(z)\big)\right)\right] \tag{6}$$

where x is real data, z is random noise, $G(z)$ is generated data, and $D(x)$ and $D(G(z))$ are the discriminator's outputs for real and generated data, respectively.

5.6. Vision Transformer (ViT)

The Vision Transformer (ViT) [35] is an adaptation of the transformer architecture [36], originally designed for natural language processing tasks for image recognition (**Figure 1.10**). ViT applies the transformer architecture to sequences of image patches, demonstrating competitive performance with traditional CNNs on image classification tasks. The critical components of ViT are as follows:

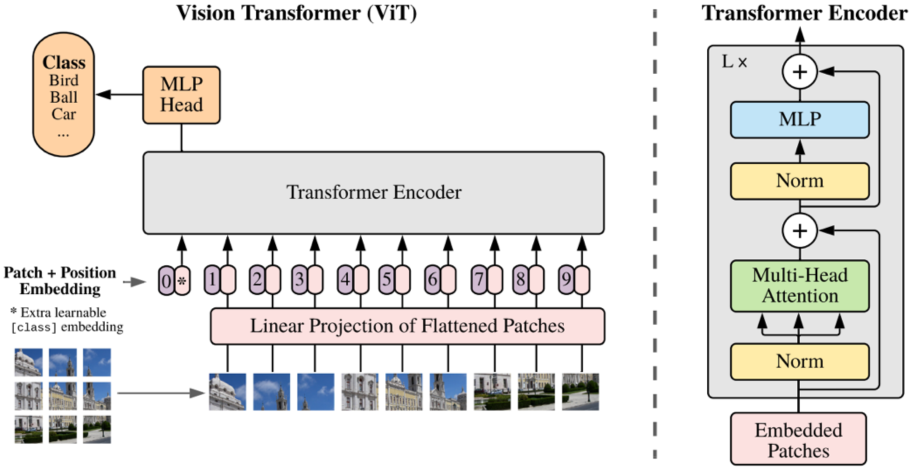


Figure 1.10 ViT Model Overview [35].

Image Patches: Instead of processing an entire image simultaneously, the ViT divides the image into fixed-size patches (e.g., 16 × 16 pixels). Each image is split into N non-overlapping patches, which are then flattened into vectors.

Patch Embedding: Each flattened patch is linearly embedded into a fixed-dimension vector using a trainable linear projection. These embedded vectors are analogous to word embeddings in original transformers in natural language processing.

Positional Encoding: Since transformers are permutation-invariant and do not inherently capture the spatial structure of images, positional encodings are added to the patch embeddings to retain information about the relative positions of patches. Positional encodings are learnable parameters added to the embeddings to provide spatial information.

Transformer Encoder: The sequence of patch embeddings with positional encodings is fed into a standard transformer encoder. The encoder consists of multiple layers of multi-head self-attention mechanisms and feed-forward neural networks. Each layer has layer normalization and residual connections, similar to the original transformer.

Classification Token: A unique learnable classification token (CLS token) is prepended to the embedded patch sequence. The output corresponding to this token from the final transformer layer is used as the image representation for classification.

Final Classification: The output of the CLS token is passed through a fully connected layer followed by a softmax activation to produce the class probabilities.

5.7. Explainable AI (XAI) in medical imaging

XAI in medical imaging is the development and application of AI models where decisions and processes are explainable to and trusted by human users, specifically medical professionals [37]. It is required because many advanced models of AI, including deep learning networks, are black boxes that may make perfect predictions but offer little insight into the reasoning behind those predictions. In medical imaging, XAI is crucial for the trust and adoption of emerging AI techniques and regulatory compliance. Some standard XAI techniques are listed below:

Saliency Maps: The visual tools highlight the regions of an image that the AI model considers most important for its decision. Techniques such as Gradient-weighted Class Activation Mapping (Grad-CAM) [38] and Layer-wise Relevance Propagation (LRP) [39] are commonly used to generate these maps.

Attention: Models can be designed with attention mechanisms that indicate which parts of the input data the model focuses on. This can provide insights into how different regions of an image contribute to the final decision.

Feature importance: Techniques such as Shapley additive explanations (SHAP) [40] and Local Interpretable Model-agnostic Explanations (LIME) [41] can identify the essential features in an image that influence the predictions by the models. This is commonly found in non-convoluted modeling practices like RF and GBM.

6. Objectives and Contributions of the Research in this Dissertation

Based on the current status of cancer image analysis and AI applications in this field, one can see there are some gaps. (1) The associations between quantitative image features and clinical events are largely unknown (e.g., patient outcome, prognosis). (2) At present, the methods for synthesizing high-quality medical imaging data, particularly for 3D volumetric data, are not yet fully mature, highlighting the current limitations in the field. (3) There is room for improvement in model accuracy by integrating multi-modal data.

To fill these gaps, two objectives are set for the research in this dissertation:

Objective One: Develop feature extraction methods to improve model accuracy.

Objective Two: Develop advanced generative models to synthesize high-quality image data.

To realize my first objective, I conducted the association study using image radiomic features to connect the medication conditions of prostate cancer patients with a GBM model (Chapter 2). To improve the feature extraction methods for cancer imaging data, I also developed an algorithm to deliver a memory-efficient ViT model with CT image data to predict pancreatic cancer patient prognosis (Chapter 5).

For the second objective, I developed an encoding method to improve the generative model with GANs to synthesize high-quality medical images (Chapter 3). To extend the method to cancer imaging data, I developed a 3D UNet-based GAN architecture to synthesize high-quality pancreatic cancer tissue images on a CT modality (Chapter 4).

Finally, in Chapter 6, I addressed my future work to address this objective in the long term.

7. References

[1] R. L. Siegel, A. N. Giaquinto, and A. Jemal, “Cancer statistics, 2024,” *CA. Cancer J. Clin.*, vol. 74, no. 1, pp. 12–49, Jan. 2024, doi: 10.3322/caac.21820.

[2] H.-P. Schlemmer *et al.*, “Global Challenges for Cancer Imaging,” *J. Glob. Oncol.*, no. 4, pp. 1–10, Dec. 2018, doi: 10.1200/JGO.17.00036.

[3] W. G. Breen, M. P. Aryal, Y. Cao, and M. M. Kim, “Integrating multi-modal imaging in radiation treatments for glioblastoma,” *Neuro-Oncol.*, vol. 26, no. Supplement_1, pp. S17–S25, Mar. 2024, doi: 10.1093/neuonc/noad187.

[4] J. R. Geis *et al.*, “Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement,” *Radiology*, vol. 293, no. 2, pp. 436–440, Nov. 2019, doi: 10.1148/radiol.2019191586.

[5] A. Singh, S. Sengupta, and V. Lakshminarayanan, “Explainable Deep Learning Models in Medical Image Analysis,” *J. Imaging*, vol. 6, no. 6, p. 52, Jun. 2020, doi: 10.3390/jimaging6060052.

[6] R. Szeliski, *Computer Vision: Algorithms and Applications*. in Texts in Computer Science. Cham: Springer International Publishing, 2022. doi: 10.1007/978-3-030-34372-9.

[7] S. Wold, K. Esbensen, and P. Geladi, “Principal component analysis,” *Chemom. Intell. Lab. Syst.*, vol. 2, no. 1–3, pp. 37–52, Aug. 1987, doi: 10.1016/0169-7439(87)80084-9.

[8] D. Bank, N. Koenigstein, and R. Giryes, “Autoencoders,” 2020, *arXiv*. doi: 10.48550/ARXIV.2003.05991.

[9] P. Kumar Mallick, S. H. Ryu, S. K. Satapathy, S. Mishra, G. N. Nguyen, and P. Tiwari, “Brain MRI Image Classification for Cancer Detection Using Deep Wavelet

Autoencoder-Based Deep Neural Network," *IEEE Access*, vol. 7, pp. 46278–46287, 2019, doi: 10.1109/ACCESS.2019.2902252.

[10] A. Krishan and D. Mittal, "Effective segmentation and classification of tumor on liver MRI and CT images using multi-kernel K-means clustering," *Biomed. Eng. Biomed. Tech.*, vol. 65, no. 3, pp. 301–313, May 2020, doi: 10.1515/bmt-2018-0175.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2015, *arXiv*. doi: 10.48550/ARXIV.1506.02640.

[13] D. L. G. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Phys. Med. Biol.*, vol. 46, no. 3, pp. R1–R45, Mar. 2001, doi: 10.1088/0031-9155/46/3/201.

[14] P. Lambin *et al.*, "Radiomics: the bridge between medical imaging and personalized medicine," *Nat. Rev. Clin. Oncol.*, vol. 14, no. 12, pp. 749–762, Dec. 2017, doi: 10.1038/nrclinonc.2017.141.

[15] R. El Ayachy *et al.*, "The Role of Radiomics in Lung Cancer: From Screening to Treatment and Follow-Up," *Front. Oncol.*, vol. 11, p. 603595, May 2021, doi: 10.3389/fonc.2021.603595.

[16] Y. Huang *et al.*, "Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer," *J. Clin. Oncol.*, vol. 34, no. 18, pp. 2157–2164, Jun. 2016, doi: 10.1200/JCO.2015.65.9128.

[17] H. Nasief *et al.*, "A machine learning based delta-radiomics process for early prediction of treatment response of pancreatic cancer," *Npj Precis. Oncol.*, vol. 3, no. 1, p. 25, Oct. 2019, doi: 10.1038/s41698-019-0096-z.

[18] A. Zwanenburg *et al.*, "The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping," *Radiology*, vol. 295, no. 2, pp. 328–338, May 2020, doi: 10.1148/radiol.2020191145.

[19] G. Litjens *et al.*, "A Survey on Deep Learning in Medical Image Analysis," 2017, doi: 10.48550/ARXIV.1702.05747.

[20] N. Tajbakhsh, L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, and X. Ding, "Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation," *Med. Image Anal.*, vol. 63, p. 101693, Jul. 2020, doi: 10.1016/j.media.2020.101693.

[21] G. A. Kaissis, M. R. Makowski, D. Rückert, and R. F. Braren, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nat. Mach. Intell.*, vol. 2, no. 6, pp. 305–311, Jun. 2020, doi: 10.1038/s42256-020-0186-1.

[22] A. Traverso, L. Wee, A. Dekker, and R. Gillies, "Repeatability and Reproducibility of Radiomic Features: A Systematic Review," *Int. J. Radiat. Oncol.*, vol. 102, no. 4, pp. 1143–1158, Nov. 2018, doi: 10.1016/j.ijrobp.2018.05.053.

[23] J. E. Van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, and B. Baessler, “Radiomics in medical imaging—‘how-to’ guide and critical reflection,” *Insights Imaging*, vol. 11, no. 1, p. 91, Dec. 2020, doi: 10.1186/s13244-020-00887-2.

[24] J. Schniering *et al.*, “Computed tomography-based radiomics decodes prognostic and molecular differences in interstitial lung disease related to systemic sclerosis,” *Eur. Respir. J.*, vol. 59, no. 5, p. 2004503, May 2022, doi: 10.1183/13993003.04503-2020.

[25] T. Bismeijer *et al.*, “Radiogenomic Analysis of Breast Cancer by Linking MRI Phenotypes with Tumor Gene Expression,” *Radiology*, vol. 296, no. 2, pp. 277–287, Aug. 2020, doi: 10.1148/radiol.2020191453.

[26] J. Dheeba, N. Albert Singh, and S. Tamil Selvi, “Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach,” *J. Biomed. Inform.*, vol. 49, pp. 45–52, Jun. 2014, doi: 10.1016/j.jbi.2014.01.010.

[27] S. Long, X. Huang, Z. Chen, S. Pardhan, and D. Zheng, “Automatic Detection of Hard Exudates in Color Retinal Images Using Dynamic Threshold and SVM Classification: Algorithm Development and Evaluation,” *BioMed Res. Int.*, vol. 2019, pp. 1–13, Jan. 2019, doi: 10.1155/2019/3926930.

[28] S. Williamson, K. Vijayakumar, and V. J. Kadam, “Predicting breast cancer biopsy outcomes from BI-RADS findings using random forests with chi-square and MI features,” *Multimed. Tools Appl.*, vol. 81, no. 26, pp. 36869–36889, Nov. 2022, doi: 10.1007/s11042-021-11114-5.

[29] D. Zikic *et al.*, “Decision Forests for Tissue-Specific Segmentation of High-Grade Gliomas in Multi-channel MR,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*, vol. 7512, N. Ayache, H. Delingette, P. Golland, and K. Mori, Eds., in Lecture Notes in Computer Science, vol. 7512. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 369–376. doi: 10.1007/978-3-642-33454-2_46.

[30] D. Zuo, L. Yang, Y. Jin, H. Qi, Y. Liu, and L. Ren, “Machine learning-based models for the prediction of breast cancer recurrence risk,” *BMC Med. Inform. Decis. Mak.*, vol. 23, no. 1, p. 276, Nov. 2023, doi: 10.1186/s12911-023-02377-z.

[31] S. Bakas *et al.*, “GLISTRboost: Combining Multimodal MRI Segmentation, Registration, and Biophysical Tumor Growth Modeling with Gradient Boosting Machines for Glioma Segmentation,” *Brainlesion Glioma Mult. Scler. Stroke Trauma. Brain Inj. BrainLes Workshop*, vol. 9556, pp. 144–155, 2016, doi: 10.1007/978-3-319-30858-6_1.

[32] G. Zhu, B. Jiang, L. Tong, Y. Xie, G. Zaharchuk, and M. Wintermark, “Applications of Deep Learning to Neuro-Imaging Techniques,” *Front. Neurol.*, vol. 10, p. 869, Aug. 2019, doi: 10.3389/fneur.2019.00869.

[33] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., in Lecture Notes in Computer Science, vol. 9351. ,

Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[34] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," 2014, *arXiv*. doi: 10.48550/ARXIV.1406.2661.

[35] A. Dosovitskiy *et al.*, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," 2020, *arXiv*. doi: 10.48550/ARXIV.2010.11929.

[36] A. Vaswani *et al.*, "Attention Is All You Need," 2017, *arXiv*. doi: 10.48550/ARXIV.1706.03762.

[37] P. Bourdon, O. B. Ahmed, T. Urruty, K. Djemal, and C. Fernandez-Maloigne, "Explainable AI for Medical Imaging: Knowledge Matters," in *Multi-faceted Deep Learning*, J. Benois-Pineau and A. Zemmari, Eds., Cham: Springer International Publishing, 2021, pp. 267–292. doi: 10.1007/978-3-030-74478-6_11.

[38] 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," *Int. J. Comput. Vis.*, vol. 128, no. 2, pp. 336–359, Feb. 2020, doi: 10.1007/s11263-019-01228-7.

[39] A. Binder, G. Montavon, S. Bach, K.-R. Müller, and W. Samek, "Layer-wise Relevance Propagation for Neural Networks with Local Renormalization Layers," 2016, *arXiv*. doi: 10.48550/ARXIV.1604.00825.

[40] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," 2017, *arXiv*. doi: 10.48550/ARXIV.1705.07874.

[41] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," 2016, *arXiv*. doi: 10.48550/ARXIV.1602.04938.

CHAPTER 2. ASSOCIATIONS BETWEEN STATIN/OMEGA3 USAGE AND MRI-BASED RADIOMICS SIGNATURES IN PROSTATE

1. Introduction

Approximately one in seven men in the US will be diagnosed with prostate cancer in their lives [1]. Prostate cancer is the most prevalent form of noncutaneous cancer and results in the second-highest number of cancer-related deaths behind only lung cancer. Since prostate cancer adversely affects so many people, it is essential to study and better understand the nature of its progression as well as patient prognosis after treatments, with the ultimate goal of improving treatment plans and their success.

Interesting correlations have been reported between prostate cancer and the use of statins and omega-3 fatty acids. Statins are a common heart medication, and omega-3 fatty acids are fish oil supplements. A study in 2018 found that there was a strong correlation between the use of statins and a reduced risk of prostate cancer metastasis, as well as a reduced risk of prostate cancer mortality [2]. Another study in 2018 focused on the effects of omega-3 fatty acids in the diet on prostate cancer tumor progression in mice and concluded that its use can inhibit the growth and development of tumor cells [3]. However, such associations have not been consistently observed among different studies, and a causative association was not fully established, nor was its mechanism uncovered [4-14]. These intriguing associations are complex and controversial, drawing active research for further elucidation. In this study, we explore the potential utility of radiomics along these efforts in analyzing the usage of these drugs among a prostate cancer patient population.

Radiomics is a new field of medicine that involves extracting large amounts of quantitative data from medical images such as CT scans or MRIs [15,16]. The data extracted are termed radiomic features, and they can be anything from the intensity of pixels in an image to the shape and texture of regions of interest (ROIs). These features can also be extracted from derived images after applying image processing filters. Radiomics, therefore, provides additional data that are often not visible to the naked eye. Furthermore, using modern big data analytics, these large amounts of quantitative data can be studied across many patients to provide a new dimension of epidemiological data that relates to medical imaging phenotypes and the underlying biological characteristics.

In prostate cancer, like in many other cancer sites, radiomics has found success in detecting and diagnosing tumors, characterizing index lesions, predicting tumor aggressiveness, evaluating treatment response and prognosis, and associating with tumor genomics [17-27]. However, to the best of our knowledge, radiomics has never been explored as a potential tool to investigate the relationship between medication exposure and prostate cancer. If radiomics signatures that are associated with the use of these drugs can be identified, such radiomics tools can be utilized to further analyze the longitudinal trends between medication exposure and prostate cancer, both during and after cancer development, based on available medical images. Using radiomics to study the correlation of statins and omega-3 in prostate cancer management adds a noninvasive, low-cost tool to assess any potential longitudinal tissue changes related to the drugs. Additionally, the presence of a defined radiomics signature that could differentiate patients who have been exposed to these medications and those who have not could help to validate the previous studies suggesting that these medications do indeed have tissue-level effects on the prostate gland as a whole as well as in prostate malignancy specifically [28,29]. The purpose of this study was to investigate the potential of radiomics as an image feature extraction technique to bridge the gap between clinical events and imaging. This is the first objective of this dissertation. By identifying potential feature patterns that could provide more understanding of the relationship between prostate cancer development and the use of these drugs, using radiomics as a tool.

2. Materials and Methods

2.1. Patient selection and clinical data collection

Male patients who developed prostate cancer and received radiation therapy treatment at our institution between 2007 and 2016 were retrospectively analyzed via electronic medical records for this study. To be included in the study, patients needed to have a definitive record of positive or negative use of statins and omega-3 supplements and a high-resolution T2-weighted MRI scan at the time of diagnosis. A total of 91 patients met the criteria and were included in the study. In addition to the diagnostic high-resolution T2-weighted MRI scan, other information collected for the study included age, ethnicity, family history, pretreatment PSA, tumor stage, NCCN risk group, and tumor grade. All data collection was approved by the IRB of our institution (Protocol: 398-17-EP). The characteristics of the studied patient cohort are summarized in **Table 2.1**. The percentage of patients with certain clinical characteristics is also listed in **Table 2.1** for each medication subgroup. Of the 91 patients, 42 used statins, and 28 used omega-3. Among them, 13 patients used both medications.

Table 2.1 Patient Demographics

	Medication Usage				All	
Parameters	Statin		Omega-3			
	Yes (%)	No (%)	Yes (%)	No (%)	N	%
Total Number of Patients	42 (46.2%)	49 (53.8%)	28 (30.8%)	63 (69.2%)	91	100%
Age at Diagnosis (median(range))	66 (54-83)y	67 (54-84)y	65 (57-82)y	68 (54–84)y	67 (54-84)y	
Race						
White	85.70%	85.70%	82.10%	87.30%	78	85.70%
Non-White	14.30%	14.30%	17.90%	12.70%	13	14.30%
Family History of Prostate Cancer						
No	78.60%	75.50%	75.00%	76.20%	70	76.90%
Yes	21.40%	24.50%	25.00%	23.80%	21	23.10%
NCCN Risk Group						
Low/Favorable Intermediate	14.30%	14.30%	17.90%	12.70%	13	14.30%
Unfavorable Intermediate	21.40%	18.40%	17.90%	20.60%	18	19.80%
High	21.40%	24.50%	21.50%	23.80%	21	23.10%
Very High	42.90%	42.80%	42.90%	42.90%	39	42.90%
Gleason Grade, Primary						
3 (4+3)	40.50%	38.80%	42.90%	38.10%	36	39.60%
4 (8)	54.80%	53.10%	50.00%	55.60%	49	53.80%
5 (9,10)	4.80%	8.20%	7.10%	6.30%	6	6.60%
Gleason Grade, Secondary						
3 (4+3)	23.80%	22.40%	17.90%	25.40%	21	23.10%
4 (8)	23.80%	24.50%	28.60%	22.20%	22	24.20%
5 (9,10)	52.40%	53.10%	53.60%	52.40%	48	52.80%
cT						
cT1c	26.20%	26.50%	21.40%	28.60%	24	26.40%
cT2						
cT2a	16.70%	18.40%	21.50%	16.00%	16	17.60%
cT2b	14.30%	8.20%	14.30%	9.50%	10	11.00%
cT2c	7.10%	10.20%	10.70%	7.90%	8	8.80%
cT3	21.40%	22.40%	17.90%	23.80%	20	22.00%
cT3a	14.30%	14.30%	14.30%	14.30%	13	14.30%
cT3b						
cN						
0	88.10%	85.70%	82.10%	88.90%	79	86.80%
1	11.90%	14.30%	17.90%	11.10%	12	13.20%

PSA					
Pretreatment PSA (median(range))	9.69 (1.31-52.19)	10.00 (1.39-59.99)	10.73 (2.00-52.19)	9.59(1.31-59.99)	9.69 (1.31-59.99)

2.2. ROI contouring and image processing

All of the patients received a high-resolution T2-weighted MRI scan in the pelvic region at diagnosis. The images were taken using Philips Medical Systems Achieva 1.5 T MRI scanners with a 2 mm slice thickness. Using Velocity software (Varian Medical Systems, Palo Alto, CA, USA), each MRI scan was manually contoured on two separate ROIs: the prostate gland and the peripheral zone of the prostate. The whole prostate gland is the most common region reviewed by radiologists and is also easy to contour for future studies, while the peripheral zone is where over 90% of prostate cancers develop [30]. **Figure 2.1** shows the created contours on two example patients, with the prostate gland outlined in blue and its peripheral region outlined in red.

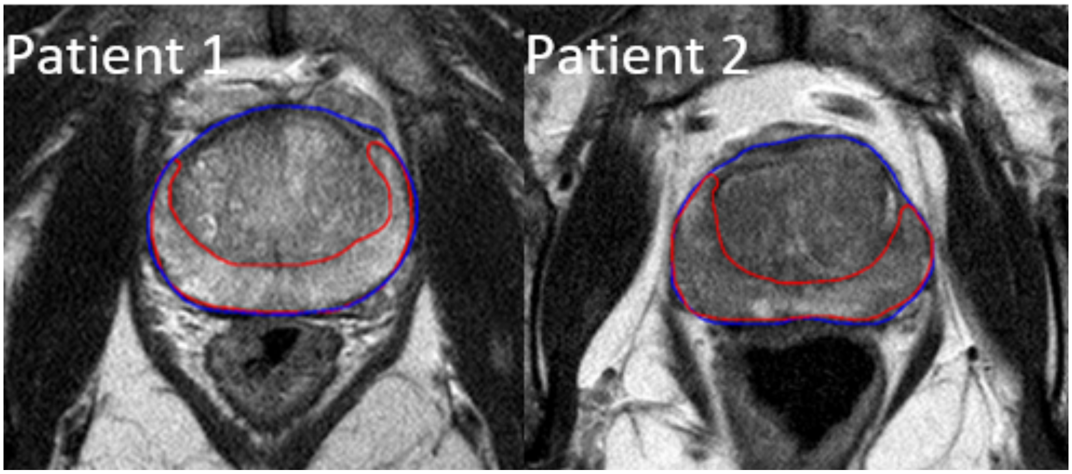


Figure 2.1 Example Contours of the Prostate ROI. Example contours of the prostate ROI (blue) and the peripheral zone ROI (red).

All of the images and segmented ROIs were then sent to 3D-Slicer v4.10 for image processing and radiomic feature extraction [31, 32]. To avoid bias field distortions and data heterogeneity bias, a bias field correction using N4 and an image normalization using histogram matching was performed on all images [33].

2.3. Radiomic feature extraction

Radiomic feature extraction was performed using the radiomics module in 3D-Slicer v4.10 [32]. Feature extraction used a resampled voxel size of 2 × 2 × 2 mm3 and a fixed bin width of 25. A kernel size of 5 was used for Laplacian of Gaussian (LoG) feature extraction. For each ROI, a total of 944 radiomic features were extracted. All of the features are defined in compliance with the feature definitions according to the Imaging Biomarker Standardization Initiative (IBSI) [34]. They are divided into the original features (107 features), LoG features (93 features), and wavelet features (744 features). The original features contain 14 shape features, 18 first-order statistical features, 14 gray-level dependence matrix features, 24 gray-level co-occurrence matrix features, 16 gray-level run-length matrix features, 16 gray-level size zone matrix features, and 5 neighboring gray-tone difference matrix features. The LoG features are calculated on the derived images by applying a LoG filter. The wavelet features are calculated on the derived images by applying either a high or low pass wavelet filter in each of the three cardinal dimensions, creating a total of 8 different filter combinations (HHH, HHL, HLL, HLH, LLL, LLH, LHL, and LHH). Because the shape features do not change on the derived images, the LoG and wavelet features do not include shape features. The full list of radiomic features is provided in **Appendix Table 7.1**.

2.4. Data analysis

Heatmaps were first generated based on each ROI to display the radiomic feature patterns for patients and their correlations with the use of statins or omega-3 supplements. Four combinations were made: Omega-3/Prostate, Omega-3/Peripheral, Statin/Prostate, and Statin/Peripheral. For each drug/ROI combination, a machine-learning model was created to select and analyze specific radiomic features that showed a correlation with drug usage. The model was trained and tested for its predictability with a workflow, depicted in **Figure 2.2**.

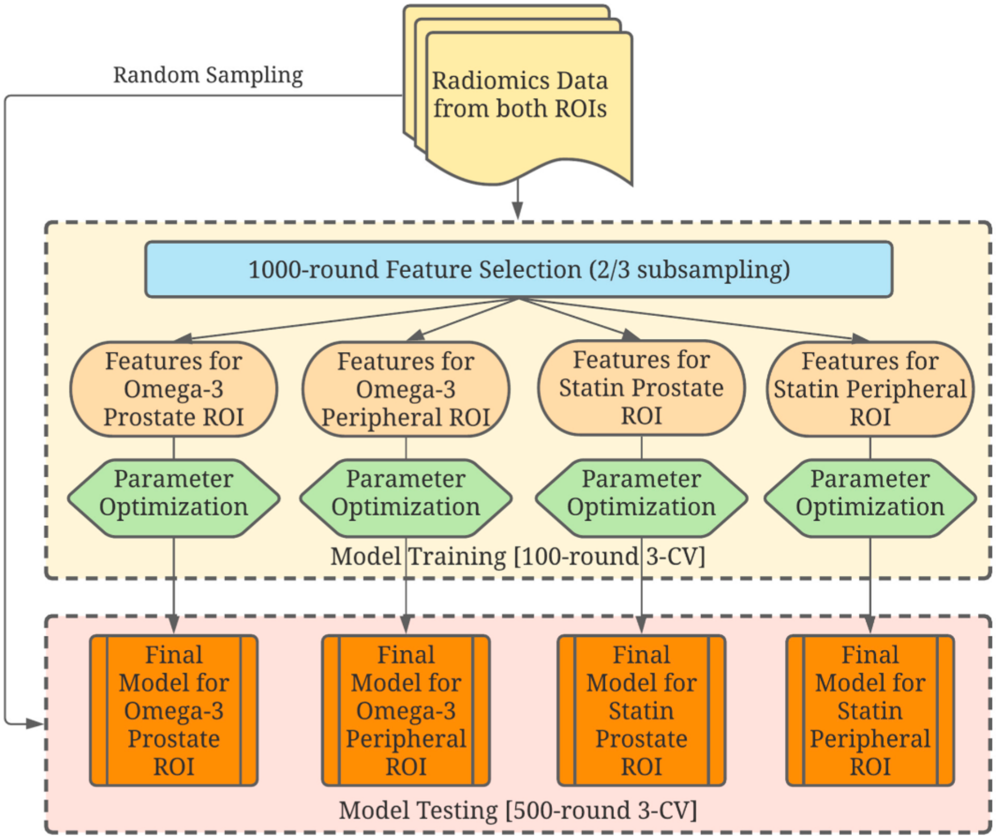


Figure 2.2 A Schematic Drawing of the Data Analysis Workflow

The four combinations of drug/ROI were studied independently in the data analysis workflow with the following steps. In the feature selection step, all patients were randomly sampled without replacement for 1000 rounds to use two-thirds of the data for feature selection and training. A univariate ANOVA analysis was first conducted on all radiomic features using the selected patients in each round. Features with a false-discovery-rate-adjusted value of $p < 0.5$ were kept for the next step. The selected features with Pearson's correlation coefficient > 0.8 were then grouped in clusters, and only the feature with the lowest p-value was kept in each cluster to reduce the dimension further [35].

To determine the learner for the next step, AdaBoost, Gradient Boosting Machine (GBM) [36], RF, Isolation Forest, and bagging classifier were tested separately with 5-fold cross-validation and parameter grid search. GBM achieved the best average AUC (**Table 2.2**).

Table 2.2 Details of Model Selection

Model	Best parameter set	AUC
RF	{'max_depth': 1, 'max_features': 0.7, 'n_estimators': 10}	0.679
GBM	{'learning_rate': 0.1, 'max_depth': 1, 'n_estimators': 10, 'subsample': 0.5}	**0.686**
Bagging	{'max_features': 0.7, 'max_samples': 0.5, 'n_estimators': 5}	0.677
AdaBoost	{'learning_rate': 1.0, 'n_estimators': 5}	0.652
Isolation Forest	{'max_features': 1.0, 'max_samples': 0.5, 'n_estimators': 5}	0.562

Consequently, GBM was selected for the final model training and the subsequent feature selection processes. A sequential floating-forward method was then utilized to select feature candidates, which gave the highest AUC value with a GBM. Only the features that could improve the AUC by more than 0.01 in each round were kept. All features were ranked by the frequency of those features selected in all 1000 rounds of resampling and the top 10 feature candidates were selected. Details of the top 10 feature candidates for the four combinations of drug/ROI are provided in **Appendix Table 7.2 – 7.5**, respectively.

Model fitting and validation were applied to GBM using the default gradient-boosting machine classifier integrated in mlr (v2.18.0) (R Foundation for Statistical Computing, Vienna, Austria) [37]. The best number of features in a model was balanced between high bias (an underfitted model) and high variance (an overfitted model). For each round of model fitting, the total search space is $8 \times 4 \times 3 \times 3 \times 3 \times 3 = 2596$, including the number of top features from 3 to 10 features (8), the number of trees (4), shrinkage (3), interaction depth (3), the minimum number of training set samples in a node to commence splitting (3), and bag fraction (3). A 500-round randomly sampled 3-fold cross-validation was applied, yielding 1500 validation tests for each drug/ROI combination. For each test, a receiver operating characteristic (ROC) analysis was conducted with the AUC and precision-recall AUC (PRAUC) values calculated. All data analysis was performed using R (version 4.0.2).

3. Results

In the heatmaps shown in **Figure 2.3**, the studied patients were clustered into similar radiomic expression patterns of the prostate gland and the peripheral region of the prostate. Information regarding statin and omega-3 use is also included. Two outlier patients were observed with extreme feature z-scores, but excluding them did not affect feature selection or model performance (data not shown), likely due to the robust resampling procedures performed in data analysis.

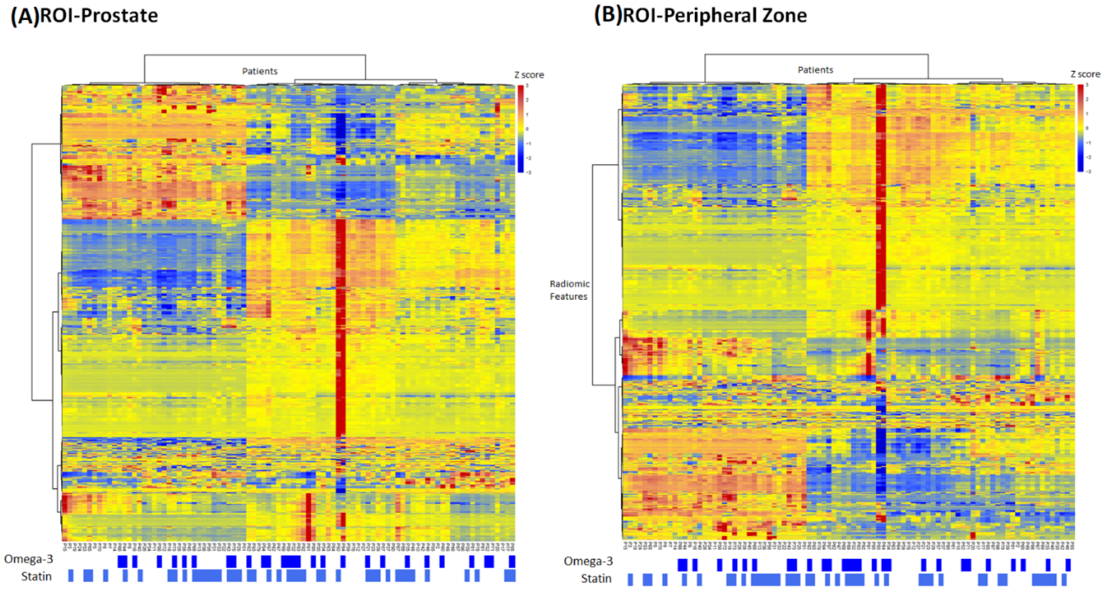


Figure 2.3 Heatmaps on Radiomic Feature Pattern. Heatmaps showing the radiomic feature pattern across the patient population with the usage of omega-3 and statin for (A) the prostate ROI and (B) the peripheral ROI.

Following the data analysis pipeline described in the Methods section, a model was then constructed and tested for each of the four drug/ROI combinations: omega-3/prostate ROI (6 features), omega-3/peripheral ROI (6 features), statin/prostate ROI (6 features), and statin/peripheral ROI (3 features). On the 1500 validation datasets from the 500-round resampled 3-fold cross-validation, the models achieved an average (standard deviation) AUCs of 0.70 (±0.08) for omega-3/prostate, 0.74 (±0.08) for omega-3/peripheral, 0.78 (±0.07) for statin/prostate, and 0.72 (±0.09) for statin/peripheral. **Figure 2.4** plots the average ROC curves from the validation analyses for these models.

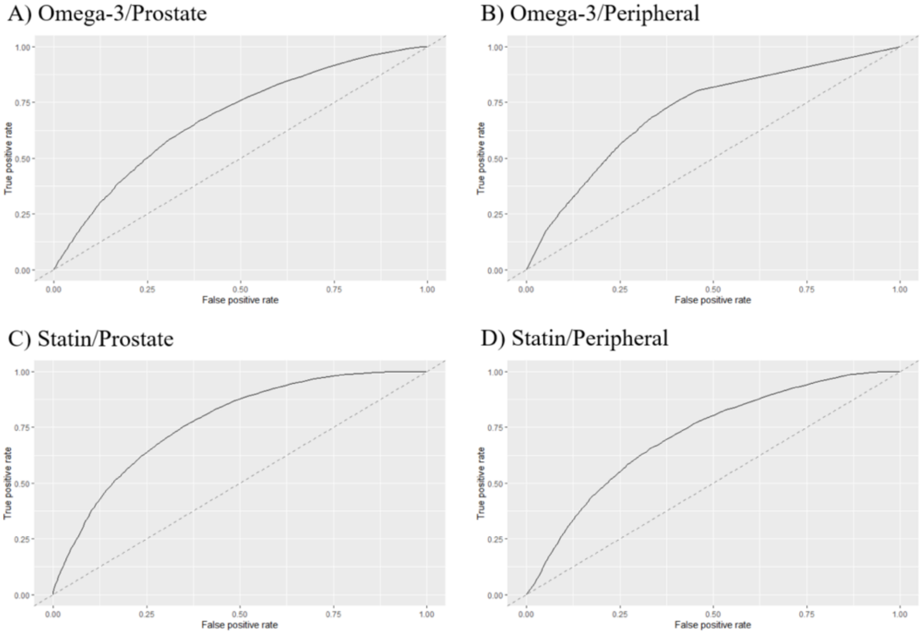


Figure 2.4 Average ROC Curves of the Radiomics Models. Average ROC curves of the radiomics model for (A) omega-3 usage prediction based on prostate ROI, (B) omega-3 usage prediction based on peripheral ROI, (C) statin usage prediction based on prostate ROI, and (D) statin usage prediction based on peripheral ROI. The ROC curves are averaged from the 1500 validation test sets.

To reveal the connection between radiomic features and drug usage, the top 6 features were listed in **Table 2.3** for each prediction. Identical radiomic features were found in two ROIs in Station prediction tasks. This finding could indicate that Statin usage may be associated with a change in the prostate's morphology and that a stable set of radiomic features could be identified. On the Omega-3 prediction tasks, no such radiomic features were found. The possible cause could be the imbalance in the Omega-3 usage data.

Table 2.3 Top 6 Feature Comparison

Prostate/Omega-3	wavelet_HLL_glszm_LargeAreaHighGrayLevelEmphasis log_sigma_5_0_mm_3D_firstorder_Mean wavelet_HLH_ngtdm_Busyness wavelet_HLH_glcm_ClusterShade wavelet_HHL_glcm_Imc1 log_sigma_5_0_mm_3D_glszm_SmallAreaLowGrayLevelEmphasis
Prostate/Statin	**wavelet_HHH_glcm_Correlation** wavelet_HHH_firstorder_Median wavelet_HLL_glcm_Correlation log_sigma_5_0_mm_3D_gldm_SmallDependenceLowGrayLevelEmphasis **wavelet_HHL_firstorder_Skewness** wavelet_HLH_glcm_Correlation
Peripheral/Omega-3	wavelet_LLL_firstorder_Kurtosis wavelet_LHH_glszm_SmallAreaLowGrayLevelEmphasis wavelet_HLL_glszm_SmallAreaEmphasis wavelet_HLH_firstorder_Mean wavelet_HLH_firstorder_Median wavelet_LHH_glszm_LowGrayLevelZoneEmphasis
Peripheral/Statin	**wavelet_HHH_glcm_Correlation** wavelet_HHL_firstorder_Median log_sigma_5_0_mm_3D_glszm_SizeZoneNonUniformityNormalized log_sigma_5_0_mm_3D_gldm_DependenceNonUniformityNormalized **wavelet_HHL_firstorder_Skewness** log_sigma_5_0_mm_3D_glszm_SmallAreaEmphasis

4. Discussion

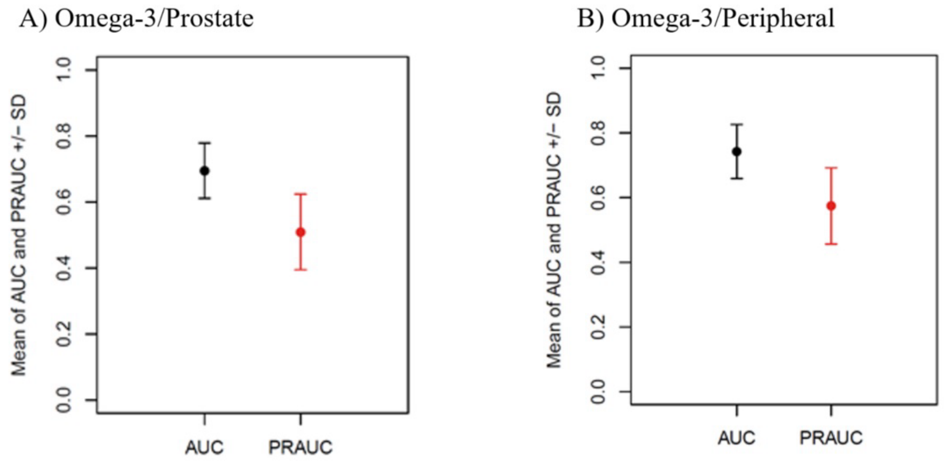
Radiomics is a novel quantitative imaging approach that involves extracting a large number of features from medical images and developing predictive models based on high-dimensional data analysis across large groups of patients. The underlying assumption is that modern high-resolution cross-sectional medical images contain information on biology and physiology that is not fully extracted or utilized by the conventional qualitative, single-patient approach. Combined with machine learning on populations of patients to uncover the otherwise hidden information from medical images, radiomics has been shown to facilitate the detection, diagnosis, and prognosis of many diseases, especially cancer [15,16,34,35].

Prostate cancer is the most prevalent non-cutaneous male cancer and is the second most lethal form of cancer for men in the United States [1]. Intra-gland and intra-tumor heterogeneity has also been well observed for prostate cancer. Based on MRI used in the routine management of prostate cancer, radiomics is well posited to study these heterogeneities as well as to assess the heterogeneity among different patients. While still in an early stage of development as a discipline, radiomics has found success in prostate cancer diagnosis, risk characterization, genomic association, and prognosis prediction, offering a non-invasive and repeatable approach in these applications [18-22,24,25,27,36]. With recent research, epidemiological, and clinical development in prostate cancer, risk stratification has become an increasingly central theme in prostate cancer management. Risk identification and prediction play a critical role in personalized clinician-patient decision-making. Radiomics, therefore, provides a potential tool for exploring the intricate interplays between medications and health supplements such as statin and omega-3 and prostate cancer that are currently active areas of risk investigations.

In this study, we explored the relationship between radiomics and the use of statins and omega-3 supplements for a cohort of 91 prostate cancer patients. This study used radiomic data based on high-resolution T2-weighted MRI, which is the most used type of medical imaging technique for prostate cancer diagnosis. Two different ROIs were investigated, including the prostate gland and the peripheral region of the prostate. Through this study, we identified specific radiomics signatures that correlated with the use of these drugs and presented machine learning models based on these signatures that were ultimately able to successfully predict the use of both medications for a given patient with prostate cancer.

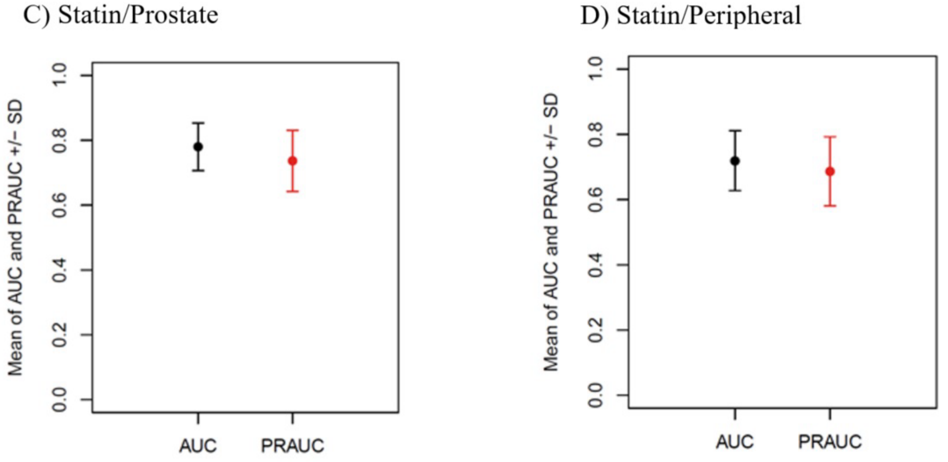
Between the two ROIs, the prostate gland showed a slightly lower AUC than the peripheral zone (0.70 vs. 0.74) for omega-3 prediction but a marginally higher AUC (0.78 vs. 0.72) for statins prediction. The finding on the identical features in the Statin usage prediction on both ROIs not only validates the previously found tissue-level effect of the medications on the prostate. However, it also preliminarily establishes radiomics signatures that could allow us to investigate the effects and dynamics of statins and omega-3 on prostate cancer development and growth further.

There are some limitations to this work. The study was a single-institution, retrospective study, with a relatively small number of patients (91 patients) meeting the selection criteria. To maximize the available data and avoid overfitting, we utilized cross-validation of a large number of random repetitions, and the number of radiomic features in the final models was also limited to 3–6 features. For omega-3, the data were more unbalanced, which could also lead to overfitting. To study its effect, we calculated both the AUC and PRAUC for each model. As expected, bigger drops were observed in the PRAUCs for the omega-3 models. **Figure 2.5** shows the validation of AUC and PRAUC achieved by each model. The selected radiomic features used in each model are also shown in **Figure 2.5**. The low performance of the Omega-3 prediction tasks may be caused by another factor. Omega-3 fatty acids are commonly found in food resources. Apart from the reported drug usage by the patients, each individual's diet could also contribute to additional variability that may not be captured in the clinical records.



Radiomic Features Included in the Model
wavelet_HLL_glszm_LargeAreaHighGrayLevelEmphasis
log_sigma_5_0_mm_3D_firstorder_Mean
wavelet_HLH_ngtdm_Busyness
wavelet_HLH_glcm_ClusterShade
wavelet_HHL_glcm_Imc1
log_sigma_5_0_mm_3D_glszm_SmallAreaLowGrayLevelEmphasis

Radiomic Features Included in the Model
wavelet_LLL_firstorder_Kurtosis
wavelet_LHH_glszm_SmallAreaLowGrayLevelEmphasis
wavelet_HLL_glszm_SmallAreaEmphasis
wavelet_HLH_firstorder_Mean
wavelet_HLH_firstorder_Median
wavelet_LHH_glszm_LowGrayLevelZoneEmphasis



Radiomic Features Included in the Model
wavelet_HHH_glcm_Correlation
wavelet_HHH_firstorder_Median
wavelet_HLL_glcm_Correlation
log_sigma_5_0_mm_3D_gldm_SmallDependenceLowGrayLevelEmphasis
wavelet_HHL_firstorder_Skewness
wavelet_HLH_glcm_Correlation

Radiomic Features Included in the Model
wavelet_HHH_glcm_Correlation
wavelet_HHL_firstorder_Median
log_sigma_5_0_mm_3D_glszm_SizeZoneNonUniformityNormalized

Figure 2.5 AUC and PRAUC Score with Top Features. AUC and PRAUC achieved on the 1500 validation sets by the radiomics model for (A) omega-3 usage prediction based on prostate ROI, (B) omega-3 usage prediction based on peripheral ROI, (C) statin usage

prediction based on prostate ROI, and (D) statin usage prediction based on peripheral ROI. The radiomic features used in each model are also listed.

To further assess the chance of overfitting despite the above mitigating measures, a control experiment with a third medication, angiotensin-converting enzyme inhibitors (ACEI) was performed. ACEI is a type of renin-angiotensin system (RAS) inhibitor and is commonly used to treat hypertension. A few studies have analyzed the association between ACEI and prostate cancer [41,42]. These studies tend to suggest that ACEI has no effect on prostate cancer development or progression. In our cohort, 35 patients (38.5%) reported using ACEIs. Applying our radiomics data analysis workflow through the ACEI usage data, we achieved an average AUC of 0.58 and an average PRAUC of 0.49 with the ACEI/prostate model (four features) and 0.54 and 0.52 with the ACEI/peripheral model (six features). Top radiomic features are provided in the **Appendix Table 7.6**. The superior model performances by the statin and omega-3 models compared with the models of the control medication, ACEI, further confirmed the validity of our radiomic findings on the two investigated medications.

Additionally, most of the patients in the dataset had their primary care physicians outside of our institution, so we could not access the initial lipid profiles to assess the reasoning behind their use of these medications or any related medical care differences that these patients may have. In our study, we individually studied the four drug/ROI combinations. The medication overlap of the 13 patients who used both medications was not considered, assuming that the possible mechanisms of prostate cancer for the two medications are independent of each other. We also did not match the other patient characteristics for each medication's positive and negative subgroups. These choices were made to maximize the sample size for the study. While biases could possibly be introduced, they are unlikely because there is currently no literature suggesting a correlation between the mechanisms of the two medications and prostate cancer, and the patient subgroups are largely balanced on other characteristics, as shown in **Table 2.1**. Furthermore, in this study, we analyzed radiomics based on imaging data only at a single time point; the time of diagnosis and the effect of varying lengths of medication usage before this time point was not investigated. While this could have acted as a confounding factor in this study, the promising findings of this preliminary study could guide future endeavors to use the radiomics tool for exploring the longitudinal tissue effects related to these medication uses in the course of prostate cancer development and progression.

5. Conclusions

For the first time, this preliminary study of two medications, statin and omega-3, in prostate cancer patients investigated the potential of radiomics in investigating drug/cancer interactions. It bridges the gap between clinical events and imaging and sets the foundation for the extraction of cancer imaging features to be utilized in downstream analytical computer vision tasks. This is an essential first step in realizing the first objective of this dissertation.

While further large-scale studies and validations are needed, encouraging findings from this study indicate that radiomics supports the tissue-level effects of these medications and that radiomics can be a valuable tool for exploring the intricate relationships between such drugs and the development and progression of prostate cancer.

This initial study revealed that the quality and quantity of image data could hinder the potential to gain insights from cancer studies. The following two chapters will focus on addressing data hungriness in cancer imaging.

6. References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70(1):7-30.

[2] Allott EH, Freeman MR, Freedland SJ. Statin therapy to improve prostate cancer outcomes: Who, when, and for how long? Eur Urol. 2018;74(6):702-703.

[3] Gevariya N, Besancon M, Robitaille K, et al. Omega-3 fatty acids decrease prostate cancer progression associated with an anti-tumor immune response in eugonadal and castrated mice. Prostate. 2019;79(1):9-20.

[4] Aucoin M, Cooley K, Knee C, et al. Fish-derived omega-3 fatty acids and prostate cancer: A systematic review. Integr Cancer Ther. 2017;16(1):32-62.

[5] Dinwiddie MT, Terry PD, Whelan J, Patzer RE. Omega-3 fatty acid consumption and prostate cancer: A review of exposure measures and results of epidemiological studies. J Am Coll Nutr. 2016;35(5):452-468.

[6] Guertin MH, Robitaille K, Pelletier JF, et al. Effects of concentrated long-chain omega-3 polyunsaturated fatty acid supplementation before radical prostatectomy on prostate cancer proliferation, inflammation, and quality of life: Study protocol for a phase IIb, randomized, double-blind, placebo-controlled trial. BMC Cancer. 2018;18(1):64-017-3979-9.

[7] Moussa H, Nguile-Makao M, Robitaille K, et al. Omega-3 fatty acids survey in men under active surveillance for prostate cancer: From intake to prostate tissue level. Nutrients. 2019;11(7):10.3390/nu11071616.

[8] Tan P, Zhang C, Wei SY, et al. Effect of statins type on incident prostate cancer risk: A meta-analysis and systematic review. Asian J Androl. 2017;19(6):666-671.

[9] Allott EH, Csizmadi I, Howard LE, et al. Statin use and longitudinal changes in prostate volume; results from the REduction by DUtasteride of prostate cancer events (REDUCE) trial. BJU Int. 2020;125(2):226-233.

[10] Carretero-Gonzalez A, Lora D, Manneh R, Lorente D, Castellano D, de Velasco G. Combination of statin/vitamin D and metastatic castration-resistant prostate cancer (CRPC): A post hoc analysis of two randomized clinical trials. Clin Transl Oncol. 2020;22(11):2126-2129.

[11] Goldberg H, Mohsin FK, Saskin R, et al. The suggested unique association between the various statin subgroups and prostate cancer. Eur Urol Focus. 2020.

[12] Meijer D, van Moorselaar RJA, Vis AN, Bijnsdorp IV. Prostate cancer development is not affected by statin use in patients with elevated PSA levels. Cancers (Basel). 2019;11(7):10.3390/cancers11070953.

[13] Moyad MA. Preventing lethal prostate cancer with diet, supplements, and rx: Heart healthy continues to be prostate healthy and "first do no harm" part II. Curr Urol Rep. 2020;21(3):15-020-0967-4.

[14] Yang H, Pang L, Hu X, et al. The effect of statins on advanced prostate cancer patients with androgen deprivation therapy or abiraterone/enzalutamide: A systematic review and meta-analysis. J Clin Pharm Ther. 2020;45(3):488-495.

[15] Lambin P, Leijenaar RTH, Deist TM, Peerlings J. Radiomics: The bridge between medical imaging and personalized medicine. Nature Reviews Clinical Oncology. 2017;14:749–762.

[16] Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.

[17] Bernatz S, Ackermann J, Mandel P, et al. Comparison of machine learning algorithms to predict clinically significant prostate cancer of the peripheral zone with multiparametric MRI using clinical assessment categories and radiomic features. Eur Radiol. 2020.

[18] Hectors SJ, Cherny M, Yadav KK, et al. Radiomics features measured with multiparametric magnetic resonance imaging predict prostate cancer aggressiveness. J Urol. 2019;202(3):498-505.

[19] Liu B, Cheng J, Guo DJ, et al. Prediction of prostate cancer aggressiveness with a combination of radiomics and machine learning-based analysis of dynamic contrast-enhanced MRI. Clin Radiol. 2019;74(11):896.e1-896.e8.

[20] Ma S, Xie H, Wang H, et al. MRI-based radiomics signature for the preoperative prediction of extracapsular extension of prostate cancer. J Magn Reson Imaging. 2019;50(6):1914-1925.

[21] McGarry SD, Bukowy JD, Iczkowski KA, et al. Gleason probability maps: A radiomics tool for mapping prostate cancer likelihood in MRI space. Tomography. 2019;5(1):127-134.

[22] Min X, Li M, Dong D, et al. Multi-parametric MRI-based radiomics signature for discriminating between clinically significant and insignificant prostate cancer: Cross-validation of a machine learning method. Eur J Radiol. 2019;115:16-21.

[23] Qi Y, Zhang S, Wei J, et al. Multiparametric MRI-based radiomics for prostate cancer screening with PSA in 4-10 ng/mL to reduce unnecessary biopsies. J Magn Reson Imaging. 2020;51(6):1890-1899.

[24] Smith CP, Czarniecki M, Mehralivand S, et al. Radiomics and radiogenomics of prostate cancer. Abdom Radiol (NY). 2019;44(6):2021-2029.

[25] Sun Y, Reynolds HM, Parameswaran B, et al. Multiparametric MRI and radiomics in prostate cancer: A review. Australas Phys Eng Sci Med. 2019;42(1):3-25.

[26] Toivonen J, Montoya Perez I, Movahedi P, et al. Radiomics and machine learning of multisequence multiparametric prostate MRI: Towards improved non-invasive prostate cancer characterization. PLoS One. 2019;14(7):e0217702.

[27] Woznicki P, Westhoff N, Huber T, et al. Multiparametric MRI for prostate cancer characterization: Combined use of radiomics model with PI-RADS and clinical parameters. Cancers (Basel). 2020;12(7):10.3390/cancers12071767.

[28] Kobayashi T, Shimizu Y, Terada N, et al. Regulation of androgen receptor transactivity and mTOR-S6 kinase pathway by rheb in prostate cancer cell proliferation. Prostate. 2010;70(8):866-874.

[29] Muller JM, Metzger E, Greschik H, et al. The transcriptional coactivator FHL2 transmits rho signals from the cell membrane into the nucleus. EMBO J. 2002;21(4):736-748.

[30] Alizadeh, M.; Alizadeh, S. Survey of clinical and pathological characteristics and outcomes of patients with prostate cancer. *Glob. J. Health Sci.* 2014, *6*, 49–57.

[31] Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D slicer as an image computing platform for the quantitative imaging network. Magn Reson Imaging. 2012;30(9):1323-1341.

[32] Van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Res. 2017;77(21):e104-e107.

[33] Tustison NJ, Avants BB, Cook PA, et al. N4ITK: Improved N3 bias correction. IEEE Trans Med Imaging. 2010;29(6):1310-1320.

[34] Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. . 2017.

[35] Karl, P. Notes on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* 1895, *58*, 240–242.

[36] Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* 2001, *29*, 1189–1232.

[37] Bischl, B.; Lang, M.; Kotthoff, L.; Schiffner, J.; Richter, J.; Studerus, E.; Casalicchio, G.; Jones, Z.M. Machine learning in R. *J. Mach. Learn. Res.* 2016, *17*, 5938–5942.

[38] Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images are more than pictures, they are data. Radiology. 2016;278(2):563-577.

[39] Aerts HJ. The potential of radiomic-based phenotyping in precision medicine: A review. JAMA Oncol. 2016;2(12):1636-1642.

[40] Stoyanova R, Takhar M, Tschudi Y, et al. Prostate cancer radiomics and the promise of radiogenomics. Transl Cancer Res. 2016;5(4):432-447.

[41] Mao, Y.; Xu, X.; Wang, X.; Zheng, X.; Xie, L. Is angiotensin-converting enzyme inhibitors/angiotensin receptor blockers therapy protective against prostate cancer? *Oncotarget* 2016, *7*, 6765–6773.

[42] Siltari, A.; Murtola, T.J.; Talala, K.; Taari, K.; Tammela, T.L.J.; Auvinen, A. Antihypertensive drug use and prostate cancer-specific mortality in Finnish men. *PLoS ONE* 2020, *15*, e0234269.

CHAPTER 3. TISSUE-SPECIFIC COLOR ENCODING AND GAN SYNTHESIS FOR ENHANCED MEDICAL IMAGE GENERATION

1. Introduction

Medical imaging and analysis play a critical role in the realm of healthcare, especially for cancer. They are crucial in identifying diseases early [1], allowing timely interventions, and improving patient outcomes [2]. They are employed in precise diagnoses, understanding disease characteristics, and devising treatment plans. Furthermore, image analysis guides surgical processes and monitors treatment effectiveness, minimizing invasiveness and complications [3]. Medical images are also indispensable in education, enhancing diagnostic skills and deepening our understanding of diseases [4]. They are integral to medical research, clinical trials, and the advancement of healthcare technologies [5].

With the rapid progression in computer vision and deep learning techniques, there has been significant progress in medical image analysis. However, unlike the typical computer vision tasks involving natural images, medical images present greater diversity, are more costly to obtain, and raise additional ethical concerns, particularly regarding patient privacy and consent. Consequently, constrained data availability significantly hinders medical imaging research and education [6]. The availability of accessible datasets is crucial for the advancements in the medical domain. Generative Adversarial Networks (GANs) are a common deep-learning approach for generating synthetic images. GANs involve a generator and a discriminator engaged in a game-theoretic competition to produce authentic synthetic data [7]. GANs are well-suited for generating artificial data as they can learn the data distribution and generate new samples that closely resemble the original data from a latent space.

While GANs have achieved remarkable success in synthesizing natural images, their application to synthetic medical images presents a unique challenge. Even though synthetic medical images can often achieve top-tier scores in image quality metrics such as FID (Fréchet Inception Distance), there is a critical concern related to the presence of artifacts that violate fundamental anatomical features. These anomalies, which compromise the anatomical accuracy of the generated images, can significantly limit the usability and reliability of synthetic medical imagery in crucial applications such as diagnostics, treatment planning, medical research, and education. Addressing and mitigating these artifacts is critical for harnessing the full potential of GANs in generating medically relevant imagery, especially for cancer tumor imaging.

This project presented a new strategy to enhance the synthesis of medical images using GANs. Rather than directly training GANs on traditional gray-scale medical images, I propose a tissue-specific GAN approach, which divides original gray-scale medical images into distinct tissue-specific representations. For example, tumor tissue will be distinguished from surrounding health tissues. These representations are then encoded into dedicated color channels of composite images, subsequently used to train GANs. Our premise is that this approach allows GANs to learn and replicate the specific characteristics of various tissues and their relationship more effectively. This tissue-specific strategy is particularly useful for cancer image synthesis, as it can enhance the accuracy of the image by emphasizing the tumor tissue. Our experimental investigations have shown that this method leverages the capabilities of GANs to generate synthetic images that can more accurately capture the subtle features of medical tissue.

In recent years, significant progress has been made in optimizing GANs; there are variants of GANs that achieve significant improvements on general image synthesis tasks by varying the objective of discriminator [8], generator [9], [10] or architecture [11] from the vanilla GAN [7]. In the medical imaging domain, modalities like MRI, CT scan, ultrasound, and radiography have different natures in image acquisition; it is necessary to reevaluate the usability or potential of the deep generative models, which are usually developed with natural image datasets.

Several works have demonstrated the generation of realistic synthetic medical imaging data. In brain MRI data generation, Shin et al. used a pix2pix conditional GAN [12] to generate brain tumor MRI images with segmentation masks [13]. The application of a PGGAN-based model was also explored in generating brain MRI images [14]. In lung cancer nodule generation, a DCGAN-based method demonstrated convincing results that were positively received by radiologists [15], [16]. Subsequently, a later model was developed to generate de-identified public radiography datasets [17]. D2FE-GAN [18] utilized decoupled dual feature representations to synthesize cross-modality MRI. FedMed-GAN [19] enhanced unsupervised cross-modality synthesis of brain images in a federated manner. 3DGAUnet [20] incorporated a 3D U-Net architecture into the generator, improving 3D shape and texture synthesis, particularly for pancreatic ductal adenocarcinoma (PDAC) tumors and pancreatic tissues.

StyleGAN [21] effectively managed features at diverse scales by incorporating modified latent vectors across various network layers at different resolutions. This approach significantly improved stability, visual fidelity, and maneuvering capabilities. Its successor, styleGAN2-ada, exhibited superior performance, especially when dealing with relatively small training datasets [22].

Despite the visual realism and high scores in image quality metrics, such as FID, synthetic medical images face challenges due to artifacts that deviate from fundamental anatomical features [23]. There's still a considerable way to go before artificially synthesized medical imaging data can be clinically usable. More than that, cancer imaging, especially 3D imaging synthesis, is an even further goal to fetch, with its heterogeneity and complex tissue composition. To achieve great things, start with small steps. This project sets the fundamentals for developing a successful, high-quality 3D cancer image generative model. Firstly, exploring a lower dimension domain, 2D X-ray images, could dramatically reduce computing needs and unforeseen technical challenges. Secondly, the experiences learned from training in 2D medical imaging data could be transferred to the development of 3D cancer imaging models. Hence, this project explores possible approaches that generate high-quality medical imaging in 2D, prioritizing anatomical integrity without drastically increasing computing overhead. This chapter proposes incorporating tissue-specific channels as an encoding mechanism for tissue-type details in the original training data as the preliminary attempt to Objective Two of this dissertation.

2. Methods

The traditional method involves training a GAN model directly on the original medical images and using the trained GAN to generate synthetic images, where both the training and synthetic images are gray-scale medical images, as depicted in **Figure 3.1(a)**. In contrast, our approach enables precise manipulation of tissue-specific image attributes within the original gray-scale medical images, facilitating the stable construction of realistic images with the desired tissue characteristics, as illustrated in **Figure 3.1(c)**. Specifically, our approach can be summarized into the following key steps:

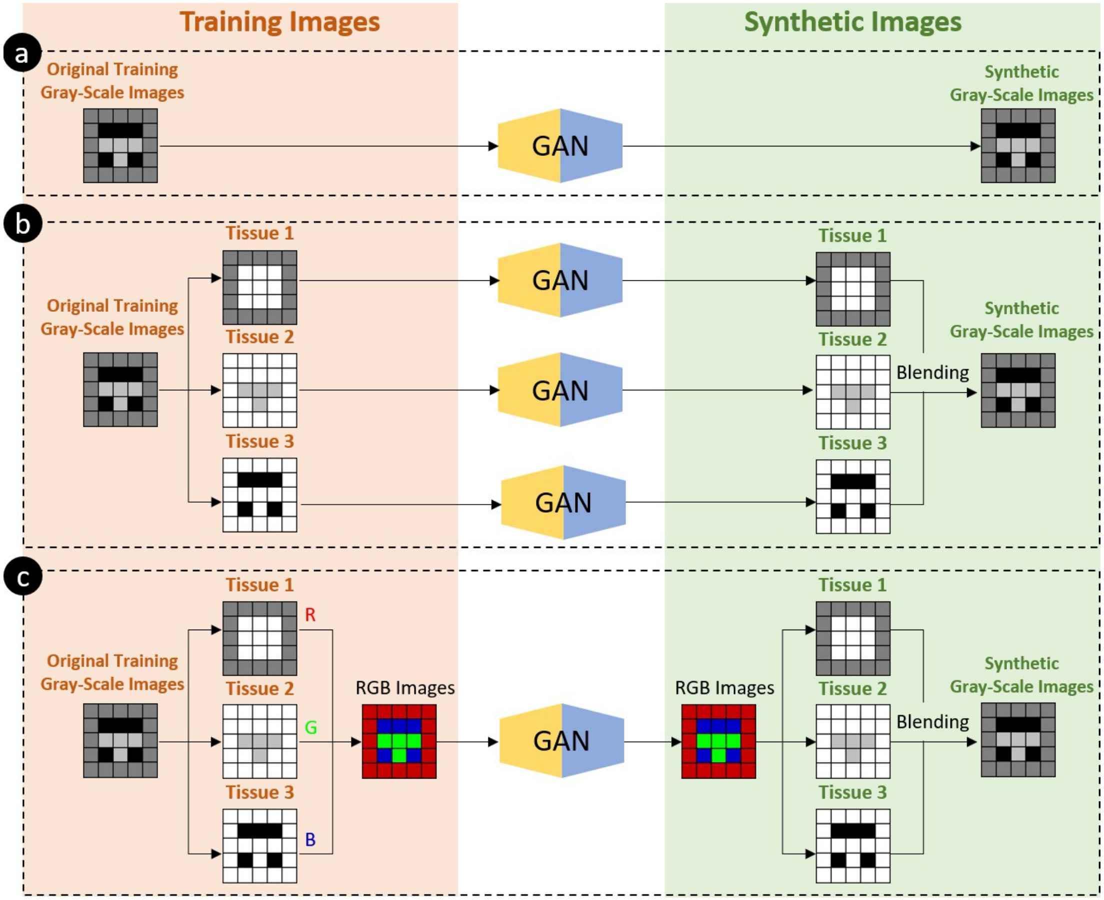


Figure 3.1 A Schematic Comparison of GANs. Comparison of (a) a traditional GAN approach, (b) an alternative non-color encoding approach, and (c) our tissue-specific color encoding and GAN Approach.

2.1. Data preprocessing

Starting with a collection of grayscale medical images, for each image, different tissues or structures are segmented and isolated, each with its unique grayscale range. Separate grayscale images for these tissues are then generated based on their ranges.

2.2. Tissue-specific color encoding

Color images were created with red, green, and blue channels. Each channel is reserved for encoding one of the gray-scale tissue images. The gray-scale images of different tissues are encoded into their respective channels, creating a multi-channel color image.

2.3. GAN training

The multi-channel color images are used as the training dataset for a GAN model. The GAN is trained to generate synthetic color images that capture the diversity of tissue features.

2.4. Post-processing

From the trained GAN model output, these gray-scale images represent the individual tissues or structures. These gray-scale images were blended to create a final synthetic gray-scale medical image and capture tissue features.

My intuition is that the effectiveness of our approach can be attributed to tissue-specific color encoding, which not only enhances the differentiation of various tissues but also improves the depiction of their spatial relationships within medical images. This approach enhances the representation of tissue features and generates more realistic synthetic medical images. An empirical study has been conducted to verify our hypothesis.

3. Experimental Design

I tested our approach using high-resolution breast coronal images. Our method enabled the generative model to create realistic breast coronal slices with the desired ratio of fatty tissue to glandular tissue. I evaluated the quality of the generated images using both qualitative assessment and quantitative metrics, comparing them to alternative approaches.

3.1. Dataset

Our model is trained on the 2023 AAPM Grand Challenge on Deep Generative Modeling for Learning Medical Image Statistics dataset [24]. We adopted this well-established image benchmark for training and testing, despite the absence of cancer tumor images, because it can provide an accurate evaluation of our methodology. The dataset contains 108,530 8-bit images with a size of 512 pixels × 512 pixels. Images are initially from VICTRE breast phantom creation software that emulates coronal slices from anthropomorphic breast phantoms. The data comprises four breast tissue composition categories: extremely dense, heterogeneously dense, fibrous and glandular density, and almost entirely fatty. Breast density is classified according to the BI-RADS system [25]. **Figure 3.2** shows two images of coronal slices from this dataset.

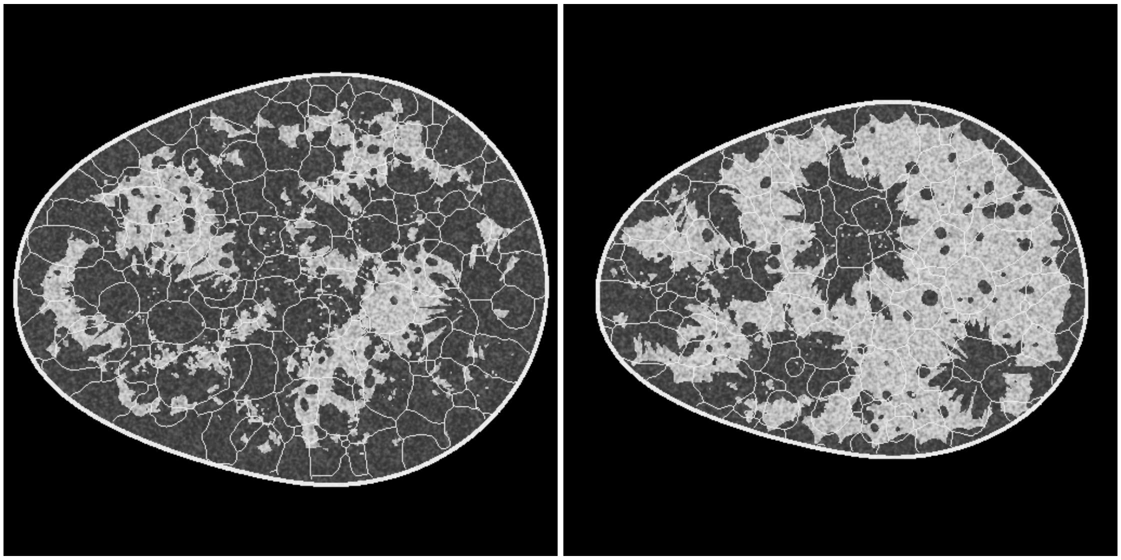


Figure 3.2 Ground Truth Image Slices. Two Ground Truth Breast Coronal Slices From the Dataset Used in Our Experiment.

3.2. Tissue-specific channels

The original gray-scale image was converted with a size of 512 pixels × 512 pixels into three channels based on its tissue-specific intensity range. The fat tissue channel includes all pixels in the range of [45, 120), the glandular tissue channel consists of all pixels in the range of [120, 226), and the dense tissue channel includes all pixels in the range of [226, 255). For each original image, a new color image with a size of 512 pixels × 512 pixels is created, and the fat, glandular, and dense tissue channels are encoded into the red, green, and blue channels of the color image, respectively. The threshold of the range is based on X-ray attenuation coefficients assigned to the various tissues on existing breast computed tomography. **Figure 3.3** shows examples of fat, glandular, and dense tissues separated from two ground truth coronal slices.

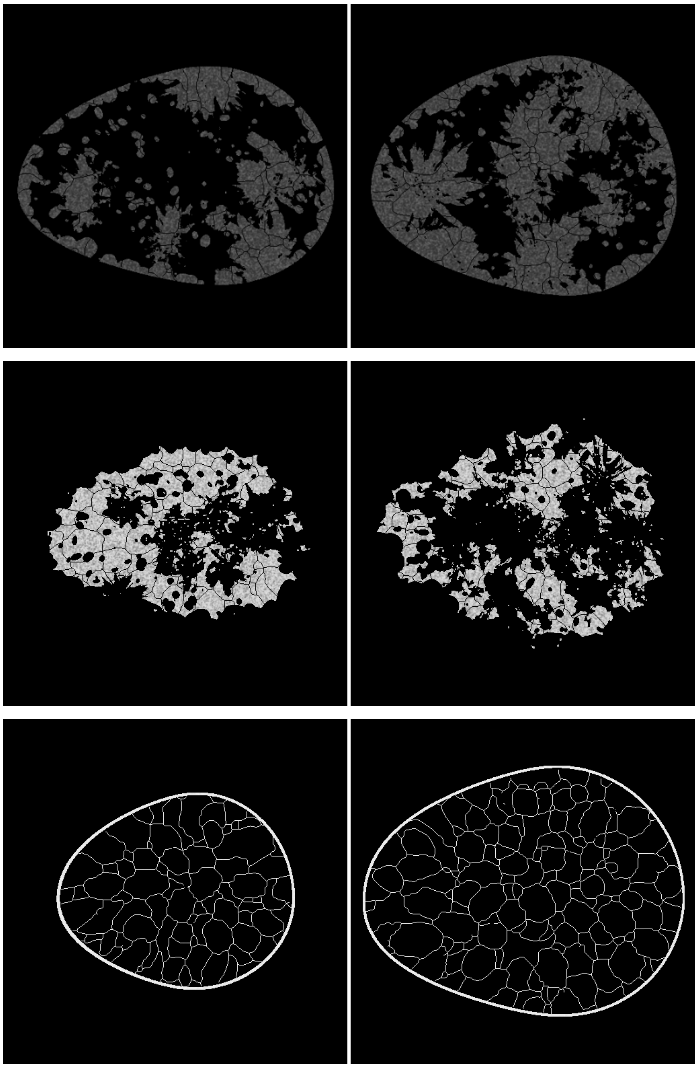


Figure 3.3 Tissue Separation. Examples of Fat, Glandular, and Dense Tissues (from the top to bottom rows) Separated From Two Ground Truth Breast Coronal Slices.

3.3. Generative model

To counter the constraint of the limited amount of data, augmentations are widely used; our generative model is based on the StyleGAN2-ada [22] since it has been reported to have top performance when dealing with limited data with a controllable augmentation pipeline. The training set has a total number of 108,530 images. The parameters are as follows: training duration = 25,000 kimg, R1-gamma = 6.6, batch size

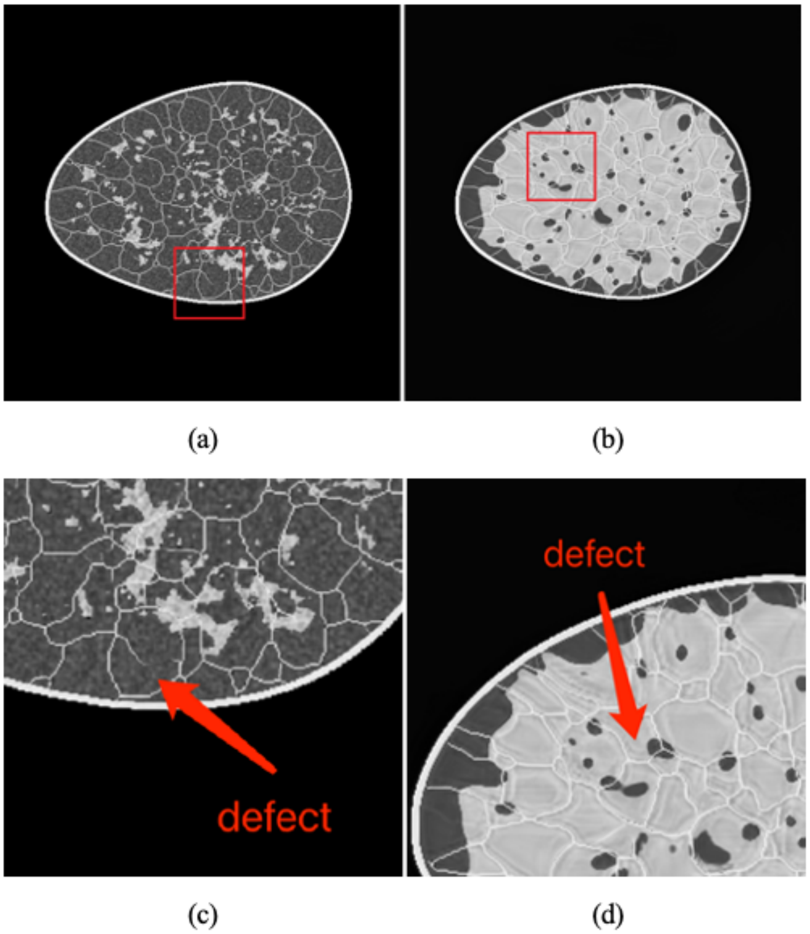
= 8, learning rate = 0.0025. The augmentation factor = 0.6, with blit, geom, color, filter, noise, and cutout image augmentation available. The best model achieved at 7600 kimg. 10,000 images were generated with the network weights for further tests. The original output of the images has three channels; each channel goes through a background noise reduction pipeline. Low-value pixels were removed, and the sum of three matrices was used to generate final gray-scale images.

3.4. Alternative non-color encoding method

In addition to the traditional GAN approach illustrated in **Figure 3.1 (a)**, we have developed an alternative method, as depicted in **Figure 3.1 (b)**, which does not utilize color encoding. In this alternative approach, tissues are segregated into distinct gray-scale images based on their unique intensity ranges, similar to our method. However, in contrast to our approach, these individual gray-scale tissue images are utilized to train separate GAN models. Subsequently, these GANs are applied to generate synthetic images for the respective tissues. The synthetic images of the individual tissues are then blended to produce a final image.

4. Results

Figure 3.4 shows two results produced by applying the traditional GAN approach, as depicted in **Figure 3.1 (a)**, with the breast images serving as the training dataset. Generally, the conventional GAN technique can generate visually pleasing results, as shown in **Figure 3.4 (a)** and **(b)**, that closely resemble ground truth images, e.g., **Figure 3.2**. However, upon closer examination, we can observe that these images have certain imperfections, failing to retain subtle tissue characteristics. For instance, when we zoom in, we notice discrepancies such as the discontinuity within the dense tissue (**Figure 3.4 (c)**) and the presence of shadows within the glandular tissue (**Figure 3.4 (d)**). These anomalies do not correspond with the expected anatomical features of these tissues.



(a) (b)

(c) (d)

Figure 3.4 Defects of the Traditional Approach. Example results using the traditional GAN approach (top row) and their defects detected upon closer examination (bottom row).

Figure 3.5 shows examples of synthetic images of fatty tissue, glandular tissue, and dense tissue that were generated by the alternative non-color encoding method. We can see that these results exhibit a remarkable similarity to ground truth tissue images, such as the ones shown in **Figure 3.3**. Furthermore, this approach could reduce the imperfections the conventional GAN method produces. We hypothesize that this is because of the means of isolating and employing individual tissues for training specific GANs, yielding better synthetic outcomes for each tissue type. The rationale is that this reduced the complexity of each image by limiting the presented tissue type to only one. However, more than an overlay of these tissue images is required to produce high-quality final synthesized results. As revealed in **Figure 3.6**, the overly of three image layers fails to capture the coherent relationship among these tissues, as seen in the ground truth images, such as **Figure 3.2**.

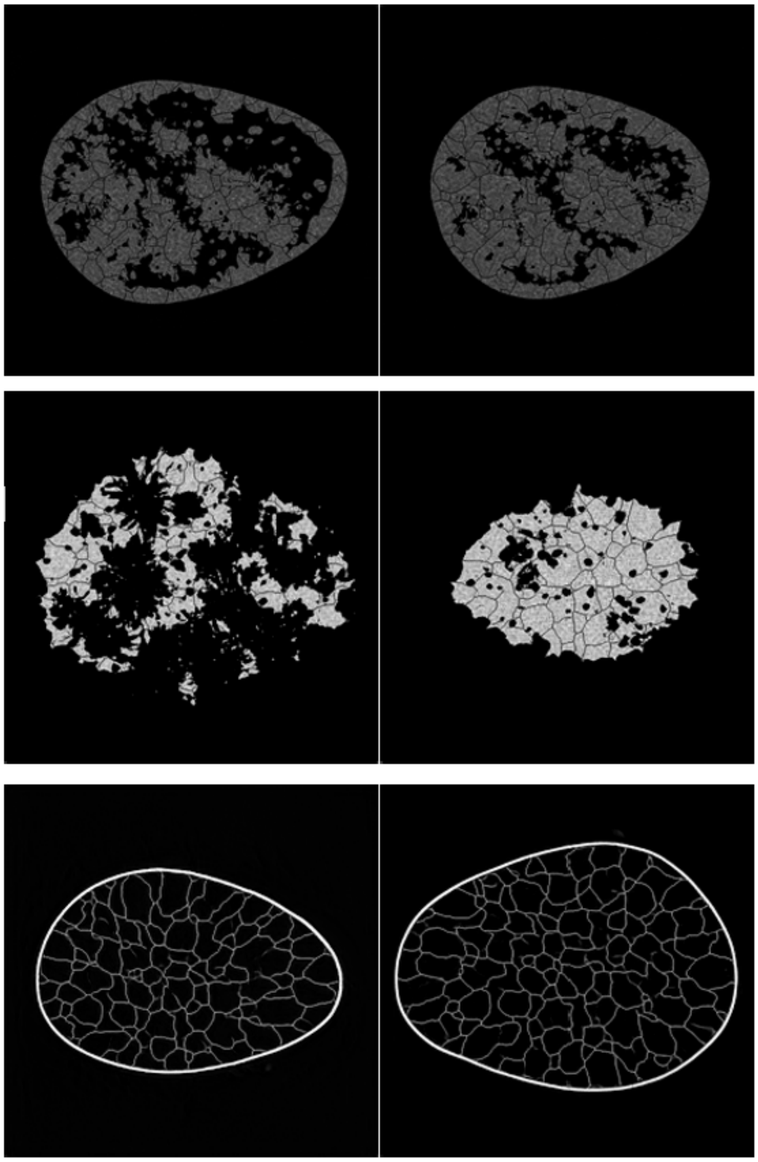


Figure 3.5 Non-color Encoding Tissue Separation. The alternative non-color encoding method generates examples of synthetic fat, glandular, and dense tissue images (from the top to bottom rows).

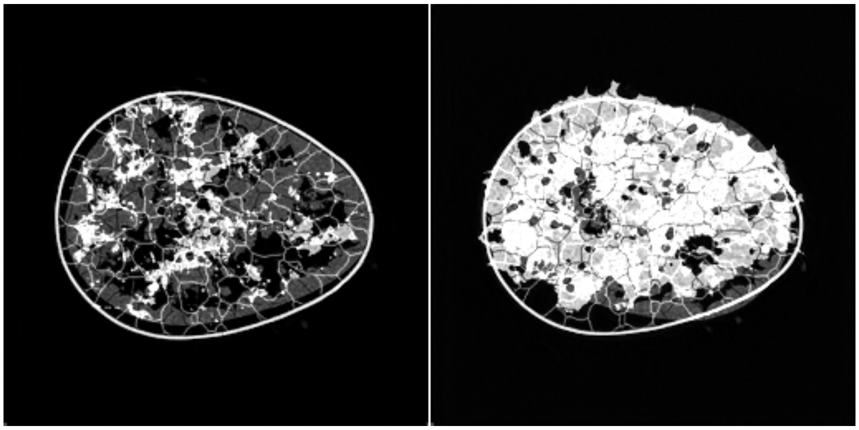


Figure 3.6 Merged Breast Coronal Slices. Final breast coronal slice results, generated by blending fatty tissue, glandular tissue, and dense tissue images synthesized using the alternative non-color encoding method, fail to capture the cohesive relationship among these tissues.

Figure 3.7 shows examples of synthetic fat, glandular, and dense tissue images generated by our method. We can see that these results also exhibit a remarkable similarity to ground truth tissue images, such as the ones shown in **Figure 3.3**, and effectively reduce the unwanted anomalies typically produced by the conventional GAN method. We postulate that this improvement is attributed to the separation of individual tissues and their encoding within distinct color channels, enabling the trained GANs to capture the unique characteristics of each tissue more precisely. In addition, a combination of these tissue images can produce high-quality final synthesized results. As revealed in **Figure 3.8**, the blended images can capture the coherent relationships among these tissues and are close to ground truth breast coronal slices, shown in **Figure 3.2**. This suggests that our method can potentially preserve the spatial connections between different tissues, and encoding each tissue image into a separate color channel can help a trained GAN model capture and retain these inter-tissue spatial relationships. Therefore, based on these qualitative experimental findings, we can see that our method can synthesize realistic fat, glandular, and dense tissue images and capture their anatomical relationship in the final photos by combining these individual tissue images.

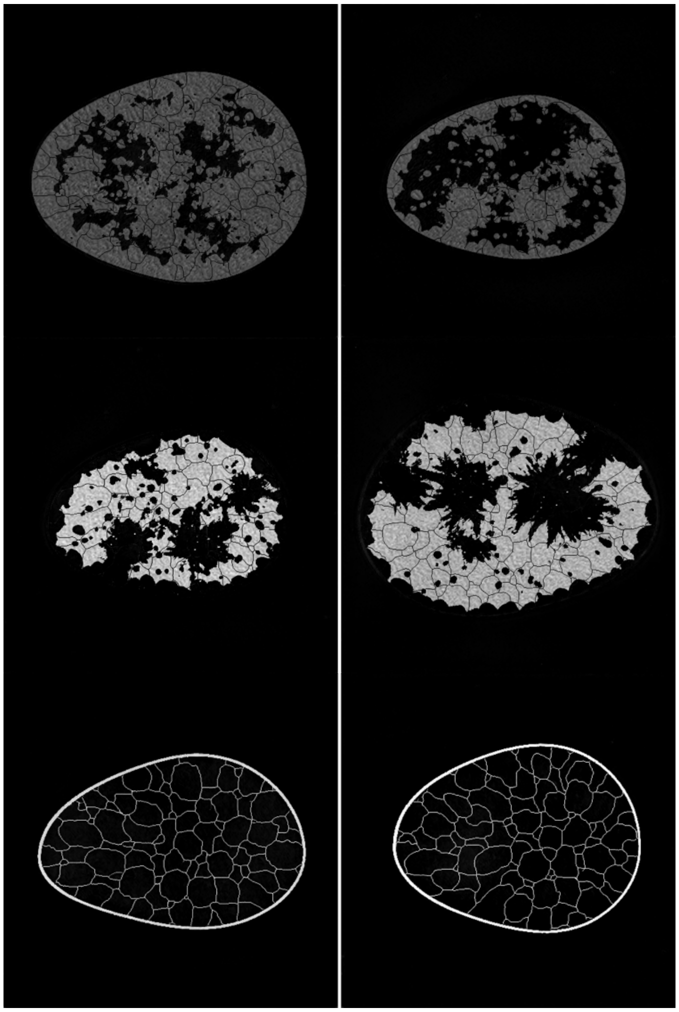


Figure 3.7 Synthetic Tissue Layers. Examples of synthetic fatty tissue, glandular tissue, and dense tissue images (from the top to bottom rows) generated by our method.

We also utilize the FID metric [26] to quantitatively evaluate the results generated by the different approaches, as summarized in **Table 3.1**. The traditional GAN approach can yield synthetic outcomes with a reasonable FID score of 13.04. However, it falls short of preserving anatomical details, as demonstrated in **Figure 3.4**.

Table 3.1 FID Scores of the Results Generated by Different Methods

Method	Traditional GAN	Non-color Encoding Approach				Our Approach			
Synthetic Tissue	Breast Coronal Slice	Fat	Glandular	Dense	Breast Coronal Slice	Fat	Glandular	Dense	Breast Coronal Slice
FID	13.04	16.28	17.92	23.14	152.75	**15.05**	**18.47**	**16.15**	**3.22**

The alternative non-encoding approach excels in faithfully synthesizing individual fat, glandular, and dense tissue images, achieved through training GANs with separate tissue representations extracted from the ground truth. However, when these synthetic tissue images are combined to produce final coronal slice results, they must attain a satisfactory FID score. The FID scores for this approach, as presented in **Table 3.1**, affirm the findings observed in **Figure 3.5** and **Figure 3.6**.

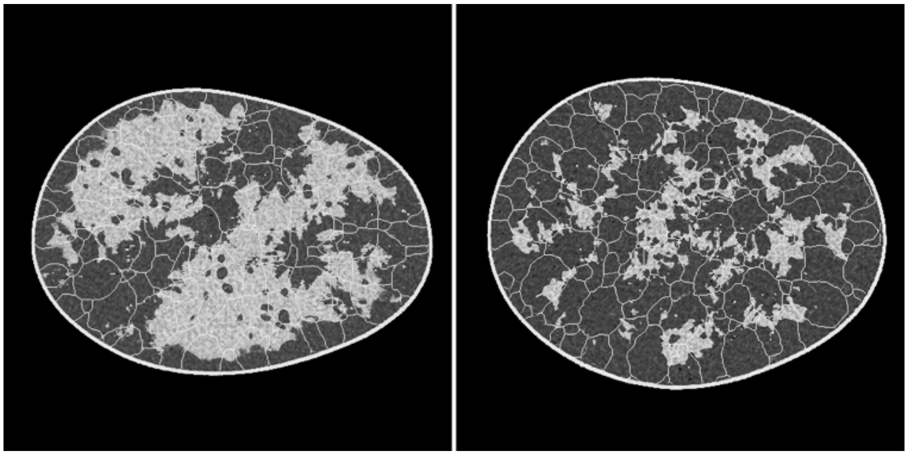


Figure 3.8 Mergerd Breast Coronal Slices from Our Method. Final breast coronal slice results, generated by blending fatty tissue, glandular tissue, and dense tissue images

synthesized using our method, can effectively capture the coherent relationships among these tissues.

In contrast, our method delivers satisfactory FID scores for individual tissue images and high-quality final breath coronal slice images. Both the qualitative (**Figure 3.7** and **Figure 3.8**) and quantitative (**Table 3.1**) findings demonstrate the effectiveness and superiority of our approach compared to the alternative methods.

My proposed encoding method verified that extending the 2D grey-scale image into 3D by storing each tissue type in one layer could improve the anatomical integrity and overall image quality. Additionally, it offers a more precise assessment of the computational demand and the manner in which the model would converge when processing 3D data. It has been a source of great inspiration for my future work on the 3D cancer imaging generative model.

5. Conclusions

We have presented a simple yet highly effective technique for medical image generation featuring tissue-specific color encoding and GAN-based synthesis. By incorporating gray-scale tissue images into the red, green, and blue channels of color images, our method significantly improves the representation of tissue characteristics and their anatomical relationships, producing synthetic medical images with enhanced realism. This innovative approach holds promise for various medical applications, offering access to high-quality, diverse, and easily interpretable synthetic medical data. Although our preliminary study focuses on encoding intensities of three tissues into the three channels of color images, it is feasible to expand our method to accommodate multiple tissue images and encode them into multiple channels within high-dimensional vector representations for GAN training. This insight has the potential to enhance the cancer image synthesis task, as the complexity was derived from the intricate tissue composition, and tissue encoding could facilitate the learning tasks.

This method can be applied to encompass a broader spectrum of tissues, especially for cancer images, to gain a more comprehensive understanding of mechanisms involved in tissue encoding. This will enable us to better capture the fundamental anatomical features and interrelationships among tissues in synthetic models, leading to more realistic medical images. In addition, we plan to engage domain experts in assessing synthesized outcomes, which will help us develop a deeper understanding of the practical implications of GAN models.

This chapter provides valuable experience for the development of more complex 3D models for cancer tumor image synthesis in the future. It serves as a simple implementation to achieve the objective two of this dissertation. This method provides a highly practical generative model encoding technique for medical imaging, bridging the gap that we described in Chapter 1.

6. References

[1] S. P. Pereira, L. Oldfield, A. Ney, P. A. Hart, M. G. Keane, S. J. Pandol, D. Li, W. Greenhalf, C. Y. Jeon, E. J. Koay, *et al.*, “Early detection of pancreatic cancer,” *The lancet Gastroenterology & hepatology*, vol. 5, no. 7, pp. 698–710, 2020.

[2] B. Qin and D. Li, “Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19,” *Sensors*, vol. 20, no. 18, p. 5236, 2020.

[3] T. M. Peters, “Image guidance for surgical procedures,” *Physics in Medicine & Biology*, vol. 51, no. 14, p. R505, 2006.

[4] B. Grignon, G. Oldrini, and F. Walter, “Teaching medical anatomy: what is the role of imaging today?” *Surgical and Radiologic Anatomy*, vol. 38, pp. 253–260, 2016.

[5] A. Hosny, C. Parmar, J. Quackenbush, L. H. Schwartz, and H. J. Aerts, “Artificial intelligence in radiology,” *Nature Reviews Cancer*, vol. 18, no. 8, pp. 500–510, 2018.

[6] H.-M. Zhang and B. Dong, “A review on deep learning in medical image reconstruction,” *Journal of the Operations Research Society of China*, vol. 8, pp. 311–340, 2020.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial networks,” *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[8] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein generative adversarial networks,” in *International conference on machine learning*. PMLR, 2017, pp. 214–223.

[9] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, “Autoencoding beyond pixels using a learned similarity metric,” in *International conference on machine learning*. PMLR, 2016, pp. 1558– 1566.

[10] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 2223–2232.

[11] T. Karras, T. Aila, S. Laine, and J. Lehtinen, “Progressive growing of gans for improved quality, stability, and variation,” *arXiv preprint arXiv:1710.10196*, 2017.

[12] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1125– 1134.

[13] H.-C. Shin, N. A. Tenenholtz, J. K. Rogers, C. G. Schwarz, M. L. Senjem, J. L. Gunter, K. P. Andriole, and M. Michalski, “Medical image synthesis for data augmentation and anonymization using generative adversarial networks,” in *Simulation and Synthesis in Medical Imaging: Third International Workshop, SASHIMI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 3*. Springer, 2018, pp. 1–11.

[14] C. Han, L. Rundo, R. Araki, Y. Furukawa, G. Mauri, H. Nakayama, and H. Hayashi, “Infinite brain MR images: PGGAN-based data augmentation for tumor detection,” in *Neural approaches to dynamics of signal exchanges*. Springer, 2019, pp. 291–303.

[15] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *arXiv preprint arXiv:1511.06434*, 2015.

[16] M. J. Chuquicusma, S. Hussein, J. Burt, and U. Bagci, “How to fool radiologists with generative adversarial networks? A visual Turing test for lung cancer diagnosis,” in 2018 IEEE 15th International Symposium on Biomedical I*maging (ISBI 2018)*. IEEE, 2018, pp. 240–244.

[17] T. Han, S. Nebelung, C. Haarburger, N. Horst, S. Reinartz, D. Merhof, Kiessling, V. Schulz, and D. Truhn, “Breaking medical data sharing boundaries by using synthesized radiographs,” *Science advances*, vol. 6, no. 49, p. eabb7973, 2020.

[18] B. Zhan, L. Zhou, Z. Li, X. Wu, Y. Pu, J. Zhou, Y. Wang, and D. Shen, "D2FE-GAN: Decoupled dual feature extraction based GAN for MRI image synthesis," *Knowledge-Based Systems*, vol. 252, p. 109362, 2022.

[19] J. Wang, G. Xie, Y. Huang, J. Lyu, F. Zheng, Y. Zheng, and Y. Jin, "FedMed-GAN: Federated domain translation on unsupervised cross-modality brain image synthesis," *Neurocomputing*, vol. 546, p. 126282, 2023.

[20] Y. Shi, H. Tang, M. J. Baine, M. A. Hollingsworth, H. Du, D. Zheng, C. Zhang, and H. Yu, "3DGAUnet: 3D generative adversarial networks with a 3D U-Net based generator to achieve the accurate and effective synthesis of clinical tumor image data for pancreatic cancer," *Cancers*, vol. 15, no. 23, 2023.

[21] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 4401– 4410.

[22] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training generative adversarial networks with limited data," *Advances in neural information processing systems*, vol. 33, pp. 12 104–12 114, 2020.

[23] V. A. Kelkar, D. S. Gotsis, F. J. Brooks, K. Prabhat, K. J. Myers, R. Zeng, and M. A. Anastasio, "Assessing the ability of generative adversarial networks to learn canonical medical image statistics," *IEEE transactions on medical imaging*, 2023.

[24] D. Gotsis, V. Kelkar, R. Deshpande, F. Brooks, P. KC, K. Myers, R. Zeng, and M. Anastasio, "Data for the 2023 AAPM grand challenge on deep generative modeling for learning medical image statistics," https://doi.org/10.13012/B2IDB-2773204 V1, 2023.

[25] C. D'Orsi, L. Bassett, S. Feig, et al., "Breast imaging reporting and data system (BI-RADS)," Breast imaging atlas, 4th ed*. American College of Radiology, Reston*, 2018.

[26] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," *Advances in neural information processing systems*, vol. 30, 2017.

CHAPTER 4. 3DGAUNET: 3D GENERATIVE ADVERSARIAL NETWORKS WITH A 3D U-NET BASED GENERATOR TO ACHIEVE THE ACCURATE AND EFFECTIVE SYNTHESIS OF CLINICAL TUMOR IMAGE DATA FOR PANCREATIC CANCER

1. Introduction

Pancreatic ductal adenocarcinoma (PDAC) represents a significant public health concern due to its delayed identification, the restricted efficacy of current chemotherapeutic treatments, and poor overall prognosis. It has the most elevated fatality rate among the primary types of solid malignancies. Despite extensive clinical and research endeavors spanning decades, the one-year survival rate is 20%. In comparison, the five-year survival rate remained in the single digits for a considerable time and only recently improved to 11% [1]. Despite the potential for a substantial increase in the 5-year relative survival rate to 42% [2] if early detection at the localized stage is achieved, there is currently a lack of definitive screening methods for reliably identifying early-stage pancreatic cancer in asymptomatic individuals.

Computed tomography (CT) is one of the primary diagnostic imaging methods. In recent years, deep-learning-based methods have increasingly been perceived as versatile applications. They can directly integrate physical and semantic details into neural network architectures [3,4,5,6,7] and are employed to solve computer vision tasks in medical imaging, such as segmentation, registration, and classification of chest X-rays and tissue histopathology images [8,9]. For example, convolutional neural networks (CNNs) have shown high feasibility in image classification tasks, in both natural and medical images, from 2D models to 3D models [10,11,12]. Some similar studies have been applied for pancreatic cancer classifiers to analyze and interpret features from medical imaging data [13,14].

While developing a deep learning model for image tasks, a substantial dataset (e.g., thousands of images) is typically needed to ensure the model converges without overfitting. Nevertheless, the availability of clinical information, particularly for PDAC, is frequently constrained by the small size of the cohorts, which presents obstacles to achieving optimal model training. Researchers have developed methods such as data augmentation, generative adversarial networks (GAN), cross-validation, and optimization approaches like sharp-aware minimization [15] to overcome the lack of training data. Generative models have demonstrated efficacy in medical image synthesis, particularly in 2D imaging modalities. Recently, researchers have developed 2D-based GAN models to generate realistic CT images of pancreatic tumors [16,17].

Nevertheless, the utilization of 3D generative models in the context of PDAC is still constrained, and directly applying existing approaches (e.g., 3D-GAN [18]) may not lead to desirable results for synthesizing three-dimensional CT image data specific to PDAC. PDAC tumors often exhibit subtle imaging features because they can be iso-attenuating or hypodense compared to the surrounding pancreatic tissue, making them difficult to distinguish visually. Additionally, PDAC tumors may lack well-defined margins, making differentiating them from normal pancreatic parenchyma challenging. Therefore, developing efficient techniques for enhancing 3D PDAC tumor datasets is crucial to facilitate the progress of deep learning models in addressing PDAC.

In this work, I developed a GAN-based tool capable of generating realistic 3D CT images depicting PDAC tumors and pancreas tissue. To overcome these challenges and make this 3D GAN model perform better, our innovation was to develop a 3D U-Net architecture for the generator to improve shape and texture learning for PDAC tumors and pancreatic tissue. The application of 3D U-Net in medical picture auto-segmentation showed appropriate and superior results. Notably, this is the first instance of its integration into GAN models. This 3D GAN model generates volumetric data of PDAC tumor tissue CT images and healthy pancreas tissue CT images separately, and a blending method was employed to create realistic final photos. Thorough examination and validation across many datasets were conducted on the developed 3D GAN model to ascertain the efficacy and applicability of the model in clinical contexts. We evaluated the effectiveness of our approach by training a 3D CNN model with synthetic image data to predict 3D tumor patches. A software package, 3DGAUnet, was developed to implement this 3D generative adversarial network with a 3D U-Net-based generator for tumor CT image synthesis. This package has the potential to be adapted to other types of solid tumors, hence making significant contributions to the field of medical imaging in terms of image processing models. This software package is available at https://github.com/yshi20/3DGAUnet.

This chapter achieved Objective Two of this dissertation. Using the insight from the last chapter, Chapter 3, my developed method provided the first-ever 3D generative model to bridge the current gap in the high-quality PDAC 3D volume synthesis task.

2. Methods

Figure 4.1 (a) illustrates the overall workflow of our proposed method. Given a set of PDAC CT images that can be acquired through different sources, we first conduct data preprocessing on these raw image data to tackle data heterogeneity and generate normalized and resampled volume data for tumor tissues and pancreas. These preprocessed datasets are then used as the training set and fed into 3DGAUnet; the new 3D GAN model developed in this work for tumor CT image synthesis. The corresponding synthetic data can be generated after the tumor and pancreas types are learned independently via 3DGAUnet.

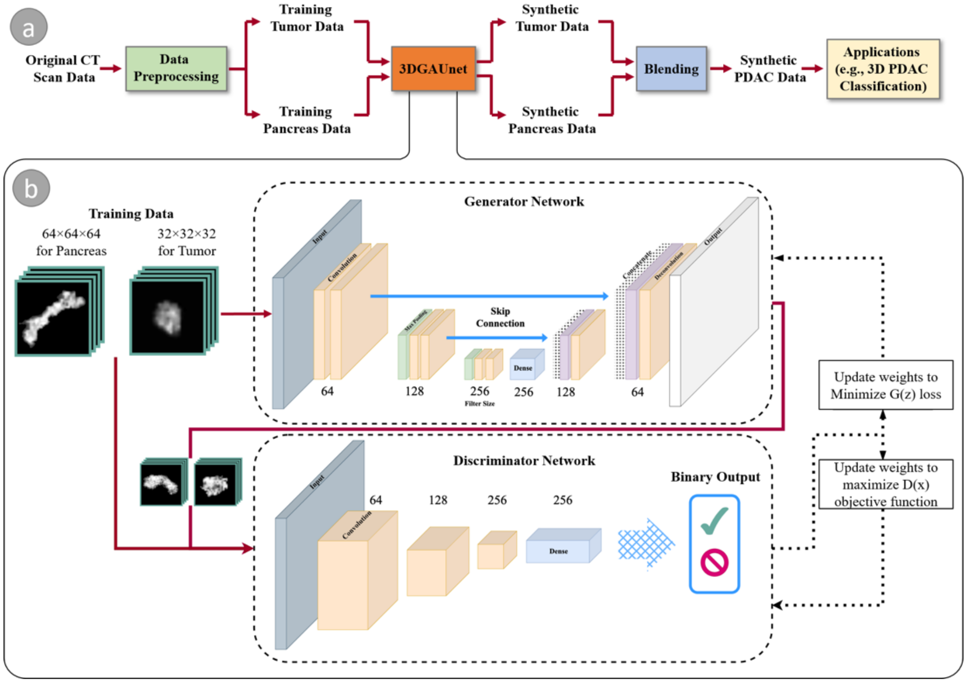


Figure 4.1 A Schematic Overview of Our Method. The workflow components (a) and the architecture of our GAN-based model (b), 3DGAUnet, consisting of a 3D U-Net-based generator network and a 3D CNN-based discriminator network to generate synthetic data.

We evaluated three blending methods to combine these synthetic tissues effectively and identified the most suitable technique for PDAC tumor CT images. Given that the pancreas is a parenchymal organ, the relative location of the tumor tissue was found to be less significant. As a result, the focus was primarily on seamlessly and realistically blending the different tissue types to ensure accurate and reliable results for diagnosing PDAC tumors in CT images.

We evaluated the usability of the synthetic data by applying it in a diagnosis task. For this purpose, we employed a 3D CNN classifier capable of taking 3D volumes as input, which improved the traditional classification tools that only use individual slices and overlook the inter-slice information.

By integrating the synthetic data, we addressed common challenges encountered in real-world scenarios, such as the small size of the dataset and imbalanced data. The addition of synthetic samples helped to improve the model's performance and mitigate issues related to imbalanced datasets.

2.1. 3D CT image data preprocessing

We used a training dataset for the GAN model, consisting of PDAC CT images from 174 patients from two sources. One was from the University of Nebraska Medical Center (UNMC) rapid autopsy program (RAP). This dataset has 71 de-identified patient data points, with the tumor contour labeled by clinical professionals (UNMC IRB PROTOCOL #127-18-EP). The other was the Medical Segmentation Decathlon pancreas training data [19], which has 103 volumetric images with a segmentation mask of tumors and pancreas. In addition, the training dataset consisted of 80 healthy pancreas CT images from the Cancer Imaging Archive (TCIA) Pancreas-CT data [20].

Normalization and resampling were essential for these raw image data, as the images were obtained from various instruments with distinct configurations. By resampling all volumetric data to 1 mm isotropic voxel spacing, each pixel in an image represents the same physical distance along each axis. For normalization, the window level threshold is an important measurement. In CT imagery, the Hounsfield unit (HU) is a dimensionless unit to measure radio density and quantify tissues within the body. It is calculated based on a linear transformation of X-ray's baseline linear attenuation coefficient, where distilled water is defined as zero HU and air is described as −1000 HU [21]. Similar HU values across different studies indicate the same type of tissue. However, calculating HU values for grayscale images with different window-level settings can lead to different visual appearances. In this study, the original images had HU values ranging from −408 to 1298. For normalization, we mapped HU values to the range of −100 to 170 for abdominal soft tissues based on the advice of an experienced radiologist.

The original image data from UNMC had metal markers that caused extremely high HU values and deflected the X-ray beam, causing the sounding tissue to have a higher HU value. To counter this defect, pixels above 200 HU were replaced with the mean HU value of the entire pancreas captured in each CT scan. Since the pancreas has an irregular shape, the tumor region and surrounding pancreas tissues that filled a cube with 64×64×64 pixels were kept as the field of interest. This made the GAN model learn the texture instead of learning the premier of the pancreas organ, which varies among different patients. After the data preprocessing step, the training volumetric data were 32×32×32 for tumor tissue and 64×64×64 pancreas cubes in grayscale.

2.2. 3DGAUnet: 3D U-Net based GAN model

We devised a 3D U-Net-based GAN model, 3DGAUnet, to synthesize 3D pancreas tumor and tissue images. **Figure 4.1(b)** shows our model architecture. At a high level, our model follows a typical GAN model that comprises two primary components, a generator *G* and a discriminator *D*. The generator creates synthetic samples. In contrast, the discriminator differentiates between synthetic and natural samples. They compete in an adversarial game to improve the generator's ability to generate genuine samples and the discriminator's ability to identify them. The aim is to produce synthetic samples that closely approximate their natural counterparts. This process can be defined as a min-max optimization task:

$$\min_G max_D L_{GAN}, \quad (7)$$

and L_{GAN} is defined as:

$$L_{GAN} = \mathbb{E}_{x\sim pdata}[\log D(x)] + \mathbb{E}_{z\sim pz}[\log(1 - D(G(z)))] \quad (8)$$

where $\mathbb{E}$ is the cross-entropy of the binary classifier of the discriminator. The task of generator *G* is to minimize the generator loss to generate synthesized images that cannot be distinguished by the discriminator *D*:

$$min_G L_{GAN} = min_G \mathbb{E}_{Z\sim pZ}[\log(1 - D(G(z)))] \quad (9)$$

The task of discriminator *D* is to separate authentic images and synthesized images better:

$$max_D L_{GAN} = max_D \{\mathbb{E}_{x \sim pdata}[\log D(x)] + \mathbb{E}_{z \sim pz}[\log(1 - D(G(z)))]\} \quad (10)$$

Our 3D GAN model employs a 3D U-Net-based structure for the 3D GAN generator and a 3D CNN-based classifier as the discriminator. The 3D U-Net structure has proven advantages for effectively capturing global and local structures, such as CT image auto-segmentation for tissues like the pancreas [22]. We developed a 3D U-Net structure for 3D image synthesis. This is the first time a 3D U-Net structure has been used for a generator in a GAN model to tackle the challenging shape and texture learning for PDAC tumors and pancreatic tissue. Each convolution layer in this model has a kernel size of 3×3×3, a stride of 2, and a ReLU activation. The skip connections allow the low-level information to be passed to the upsampling stacks to avoid the vanishing gradient problem.

The discriminator is used to identify if the input image is synthesized images from the generator's output. It has three 3D convolutional blocks, each starting with a 3D convolution with a kernel size of 2×2×2 and sides of 1. This is followed by a 3D max-pooling layer with a pool size of 2 and batch normalization. After three 3D convolutional blocks, this is flattened to a fully connected dense layer, and binary output is generated by the sigmoid function.

Our approach can be used to train 3D models of tumors and healthy pancreatic tissue separately. The training procedure concerning discriminator loss was optimized. A total of 500 3D tumor and pancreas volumes were synthesized. Inserting tumor tissue into the pancreas cube could generate a 3D volume of the tumor with surrounding pancreas tissues.

2.3. Blending to create PDAC tissues

Because the tumor and pancreas tissue volumes were generated separately, it was essential to consider how to combine the tumor and pancreas volumes to create realistic tumorous pancreas tissue. We experimented with and compared three blending methods for merging a pancreas's tumor and tissue volumes. The first method (Blend I) was a straightforward copy-and-paste operation using the tumor voxels to replace the corresponding pancreas tissue voxels. Our second and third methods (Blend II and Blend III, respectively) were inspired by DeepImageBlending, a deep learning technique that improves Poisson image blending [23]. Deep image blending is a two-stage image blending algorithm. First, it generates a seamless boundary for the source region to eliminate visible seams. Then, it refines the source region by matching styles and textures with the target image. The algorithm uses a differentiable loss function based on the Poisson equation and can handle various image styles, including stylized paintings. It achieves visually consistent blending without relying on training data. Our Blend II and Blend III methods are the first and second stages of the tool, respectively. The motivation for comparing the two stages was because, in a natural image blending task, the object should look natural and share a similar style with the background image, but this might not be true for a CT image. Unlike natural images acquired from light reflection from objects, CT images are created by recording the X-ray beam attenuation from different directions. Therefore, a presumption of a similar style may not be valid. To find the best blending method, we used visual inspection and Fréchet inception distance (FID) values to compare the three blending methods. After synthetic tumors and healthy pancreas tissues were directly output by our 3D GAN model, they were blended to generate synthetic PDAC tissues with the best blending method. A comparison of the blending methods will be provided in Section 3.2.

2.4. Evaluation of synthesized images

The performance of our developed 3DGAUnet model was evaluated both qualitatively and quantitatively. We visualized the generated volumes with 2D cross-slices and 3D volume rendering for qualitative evaluation [24]. For quantitative evaluation, we used FID values on 2D slices [25].

We propose a 3D evaluation metric, called Fréchet 3D distance (F3D), for comparing the Fréchet distance of the activation layer from a 3D CNN network with the quality of the 3D GAN model. The distance d is calculated as

$$d^2=||\mu_1-\mu_2||^2+Tr(C_1+C_2-2\sqrt{C1C2}) \tag{11}$$

where μ_1 and μ_2 are the feature-wise means of the real and synthesized images, C_1 and C_2 are the covariance matrix of the feature vectors for the real and synthesized images, and Tr is the trace linear algebra operation that is the sum of the elements along the main diagonal of the square matrix.

To calculate the μ and C, we needed the feature vector from the last pooling layer out of a pre-trained neural network. The original FID used a trained Inception V3 model [26]. Our approach, instead, used a 3D CNN with 17 layers, including four 3D convolutional blocks with a fixed random state 42, and the feature vector was the flattened layer after the last convolutional block, having a length of 512. Samples were compared in batches, and then the μ and C could be calculated with the matrix comprising each sample's feature vector.

In addition, the quality of the images generated by our developed 3DGAUnet model was evaluated using the squared maximum mean discrepancy (MMD^2), which employs kernel functions in the reproducing kernel Hilbert space to quantify the disparity between two distributions [27]. In this study, we also used pair-wise multi-scale structural similarity (MS-SSIM) to assess the diversity of the images generated by our 3DGAUnet model. MS-SSIM is a metric that quantifies the perceptual diversity of generated images by calculating the mean of MS-SSIM scores for pairs of these images [28]. This measurement allowed us to evaluate the level of variation and dissimilarity among the generated samples, providing insights into the model’s ability to produce diverse and distinct images.

2.5. 3D CNN PDAC classifier

One of our objectives in creating synthetic data was to improve the performance of PDAC tumor identification. Currently, the limited available data are an obstacle. To test our developed 3D GAN model, we built and trained a 3D CNN classifier using the synthetic data generated by our 3DGAUnet model.

We developed a 17-layer 3D CNN model [29] to test if a 3D volumetric input was healthy pancreas tissue or a tumor. The 3D CNN classifier has four 3D convoluted blocks (Conv 3D), with the first block consisting of 64 filters followed by 128, 256, and 512 filters, all with a kernel size of 3×3×3. Each Conv3D layer is followed by a max-pooling (MaxPool) layer with a stride of 2, ReLU activation, and batch normalization layer

(Batch Norm). This 3D CNN model has four Con3D-MaxPool-BatchNorm blocks and is intended to capture visual features from coarse to fine. The final output first flattens the production of the last convolutional block and passes it to a fully dense layer with 512 neurons. A dropout layer with a tunable dropout rate follows to prevent overfitting. The output is then passed to a 2-neuron dense layer with a sigmoid function for binary classification output. Because the input dimensions are 64×64×64, a relatively simple task, the architecture of the classifier was designed in a simple way to avoid the overparameterization problem, with 1,351,873 learnable parameters.

The binary classification performance was calculated from the confusion matrix. Given that TP, TN, FP, and FN correspond to a true positive, true negative, false positive, and false negative, respectively, the results were measured using precision $TP/(TP+FP)$, recall $TP/(TP+FN)$, true positive rates $TPR=TP/(TP+FN)$, and the false positive rates $FPR=FP/(FP+TN)$. The area under the curve (AUC) was calculated from the receiver operating characteristic (ROC) curve, which was plotted as true positive rates against the false positive rates under different cutoffs or as the precision against the recall.

3. Results

3.1. 3D volumetric tissue data generation

We trained our 3DGAUnet model separately using PDAC tumor and healthy pancreas data. These are referred to as the tumor and pancreas models. The tumor model was trained using PDAC data, including 174 volumetric tumor data in Nifty format. The pancreas model was trained using healthy tissue data, including 200 volumetric data in

Nifty format. Both input datasets resulted from the preprocessing steps outlined in Section 2.1.

Image augmentation, including image flipping and rotation, was performed on the training data. The augmented data for each volume were generated by rotating each volume on three axes in 12°, 24°, 36°, 48°, and 72° increments. All images for the tumor model were resampled to 1 mm isotropic resolution and trimmed to 32×32×32 size. All images for the pancreas model were resampled to 1 mm isotropic resolution and trimmed to 64×64×64 size, with pancreas tissue filling the entire cube.

The training procedure of any GAN model is inherently unstable because of the dynamic of optimizing two competing losses. For each model in this study, the training process saved the model weights every 20 epochs, and the entire model was trained for 2000 epochs. The best training duration before the model collapsed was decided by inspecting the generator loss curve and finding the epoch before the loss drastically increased. We trained our models with an NVIDIA RTX 3090 GPU. The optimal parameter set was searched within a parameter space consisting of batch size and learning rate, where the possible batch sizes included 4, 8, 16, and 32, and the possible learning rates included 0.1, 0.01, 0.001, 0.0001, and 0.00001.

The tumor and pancreas models generated 500 synthetic volumetric data separately. We first conducted a qualitative comparison between the training image sets and the synthetic image sets. We used volume rendering to visualize these datasets and inspect the 3D results. **Figure 4.2** shows examples of ground truth tumor volumes, synthetic tumors generated by the existing technique 3D-GAN [18], and synthetic tumors generated by our 3DGAUnet. We can see that when we trained our 3DGAUnet based on the group truth inputs, our model could generate synthetic tumors with realistic anatomical structure and texture and capture overall shape and details. Nonetheless, 3D-GAN produced unsuccessful data or failed to generate meaningful results in capturing the tumor's geometry. By comparing the generated pancreas volumes with authentic medical images, we can see that our 3DGAUnet could effectively synthesize a 3D pancreas to resemble actual anatomical structures. However, it was hard for 3D-GAN to generate anatomically plausible results, and a certain ambient noise was perceived in the generated volumes. We further examined the interior structures of the volumes generated by our 3DGAUnet. **Figure 4.3** shows the 2D slices of the ground truth, 3D-GAN, and our 3DGAUnet images from both tumor and pancreas models. The synthetic data produced by our 3DGAUnet model exhibited a high degree of fidelity to the ground truth in terms of internal anatomical structure and texture compared to 3D-GAN. In certain instances, 3D-GAN failed to produce meaningful outcomes.

However, we can also observe marginal defects in the pancreas generation, with tiny tissues surrounding the primary tissue generated in the center. These defects were likely due to the pancreas' irregular shape, different sizes, and direction, as well as the gradient learned from the input image batches at specific locations turning into noise.

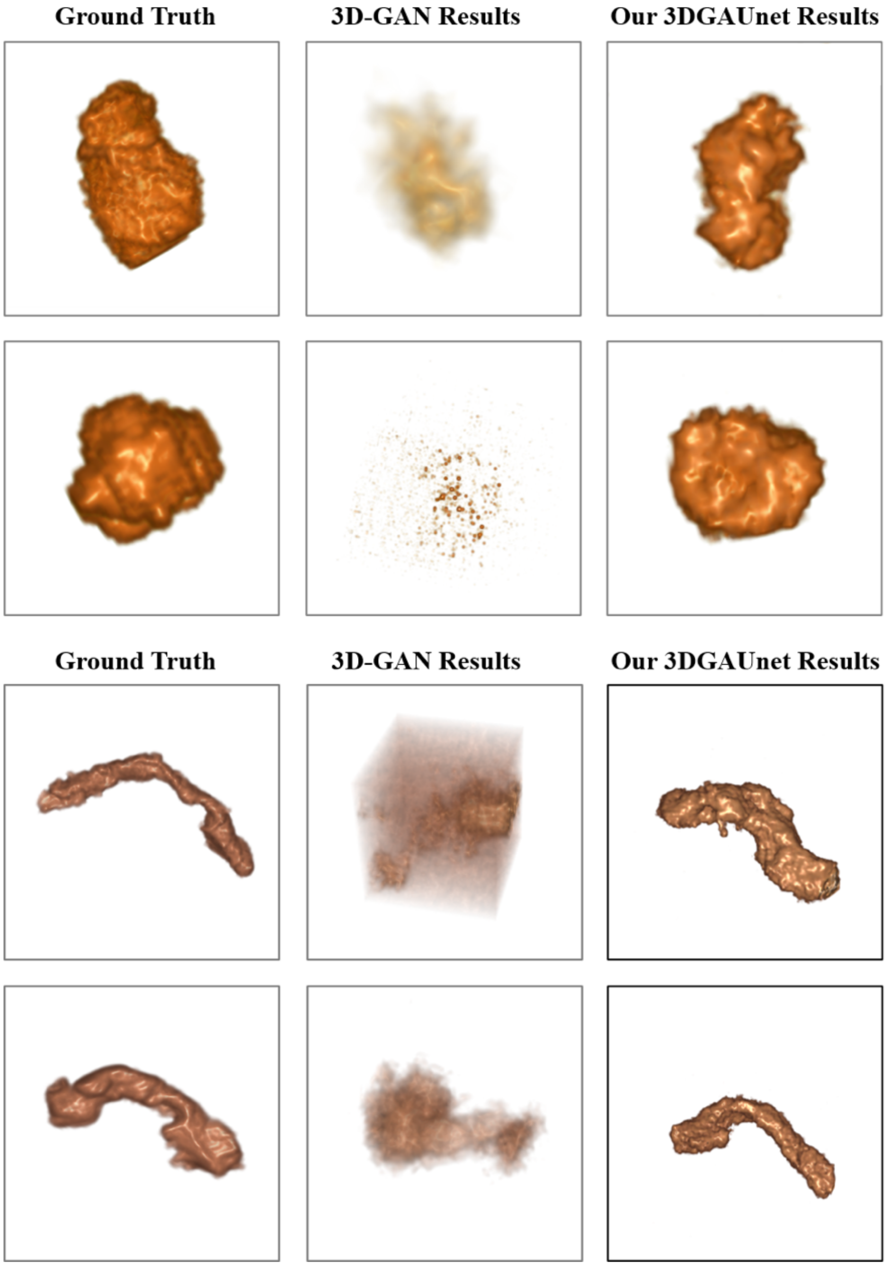


Figure 4.2 Examples of 3D Volume Data. Examples of 3D volume data of tumor (top) and pancreas (bottom) from the different methods. In each set of examples, the left, middle, and right columns correspond to ground truth data, synthetic data generated by 3D-GAN, and synthetic data generated by our 3DGAUnet, respectively. All 3D volumes are shown in volume rendering.

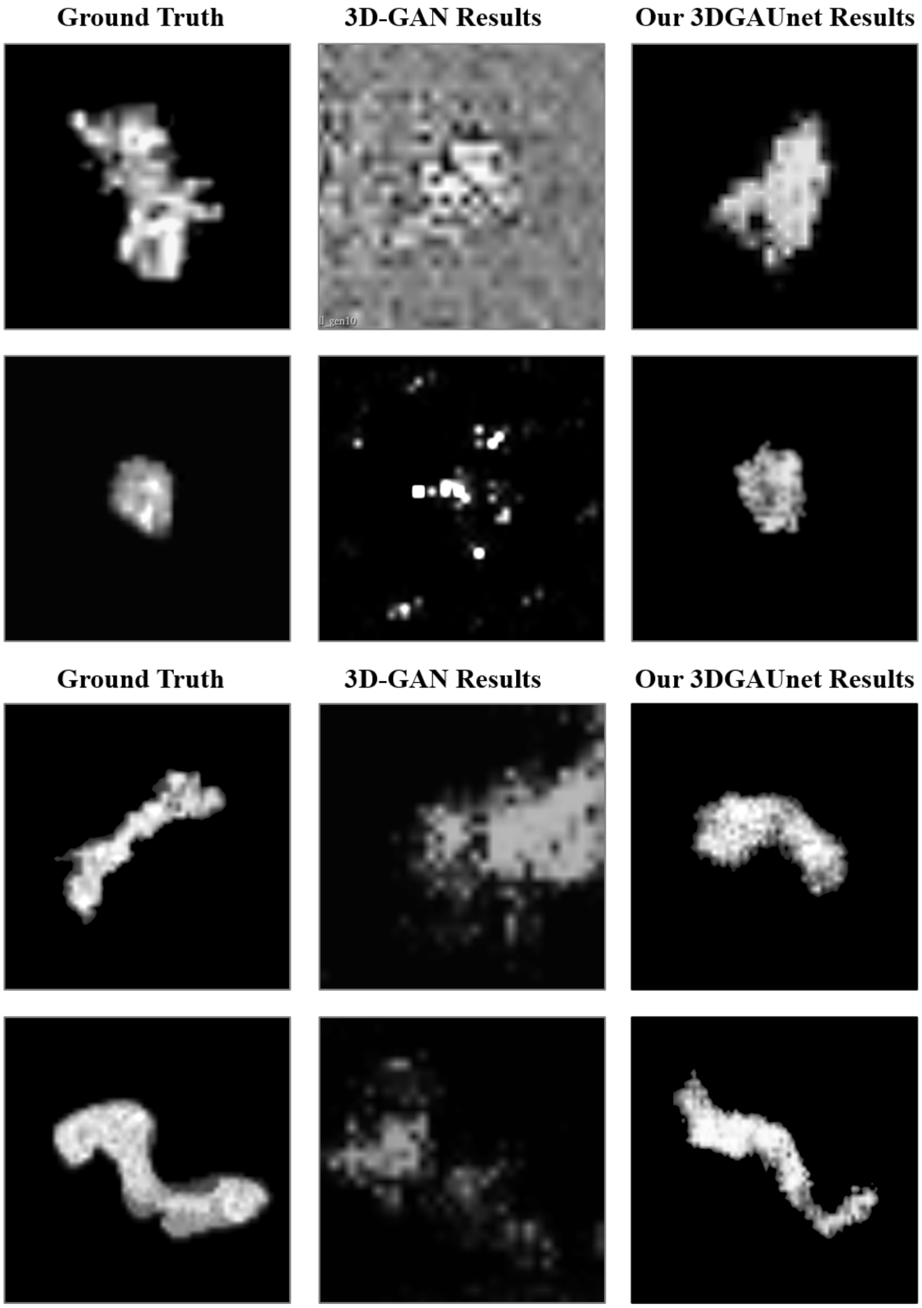


Figure 4.3 Example of 2D Slice in 3D Volume. Examples of 2D in 3D volumes of tumor (top) and pancreas (bottom) from the different resources. In each set of examples, the left, middle, and right columns correspond to ground truth data, synthetic data generated by 3D-GAN, and synthetic data generated by our 3DGAUnet, respectively.

In addition, we conducted a quantitative assessment of the outcomes produced by 3D-GAN and our 3DGAUnet models. A sample of 100 ground truth and 100 synthetic volumes was randomly selected. Specific 2D image metrics, such as slice-wise FID and PSNR, were computed to assess the quality of synthetic volumes. Besides the 2D metrics, 3D metrics such as batch-wise F3D, MMD^2, and SSIM were calculated on the randomly selected volumetric data. **Table 4.1** shows the values of 2D image metrics, slice-wise FID, and slice-wise PSNR on the sagittal (Sag), axial (Ax), and coronal (Cor) planes to estimate the quality of the synthetic volumes, where scores were calculated using the center slice from 100 synthesized volumes and 100 ground truth volumes from the tumor model and pancreas model separately. **Table 4.2** shows the values of the 3D volume metrics, batch-wise F3D, the MMD^2, and the MS-SSIM. From the results, we can observe that our 3DGAUNet outperformed 3D-GAN in all metrics, suggesting that 3DGAUNet excelled at capturing the 3D shape and texture characteristics for both tumor and pancreas compared to 3D-GAN. In addition, all the quantitative metrics in the pancreas model were better than the tumor model, especially on FID and F3D. This was probably caused by the difference in the training tasks, where the tumor model needed to learn both the texture and shape of the tumor, but the pancreas model was trained with pancreas-filled cubes to know the texture of the pancreas image mostly.

Table 4.1 Performance Based on 2D Image Quality Metrics

Tissue	Model	FID-Sag	FID-Ax	FID-Cor	PSNR-Sag	PSNR-Ax	PSNR-Cor
Tumor	3D-GAN	249.32	262.18	244.27	20.10	18.63	19.49
	3DGAUNet	198.23	202.44	188.66	16.52	17.76	17.16
Pancreas	3D-GAN	293.62	342.60	335.20	18.20	16.31	14.05
	3DGAUNet	287.75	435.72	327.41	12.73	7.21	9.42

Table 4.2 Performance Based on 3D Image Quality Metrics

Tissue	Model	F3D	MMD^2	MS-SSIM
Tumor	3DGAN	472.64	5571.90	0.86
	3DGAUNet	271.31	5327.32	0.81
Pancreas	3DGAN	889.40	8924.39	0.83
	3DGAUNet	872.33	9122.40	0.77

3.2. 3D Volumetric data blending

We selected 100 pairs of synthetic tumors and synthetic pancreas tissue volumes from the synthesized data. The paired data were used to evaluate three previously introduced blending methods. Quantitative evaluation was conducted by comparing the FID Score, while qualitative evaluation was conducted by visualizing the 2D slices. 100

random abdominal CT images were cropped into 64×64×64 cubes as a negative sample set and then used to compare the 3D metrics with the blended volumes.

Figure 4.4 shows the 2D slices of the blended images using the three different blending methods. It is evident that the direct copy-and-paste approach consistently yielded the least favorable results. This was because the tumor object was extracted with a simple pixel value threshold; therefore, the tumor tissue boundary may not have been as precise as needed. One can spot black pixels randomly appearing around the boundary of the tumor and pancreas tissue. Meanwhile, Blend III was visually closer to the ground truth tumor site and, on average, had a lower slice-wise FID score. Therefore, we employed Blend III as the blending method for the developed 3D GAN model. **Table 4.3** compares the slice-wise FID values among the three blending methods and clearly shows that Blend III achieved the best slice-wise FID values.

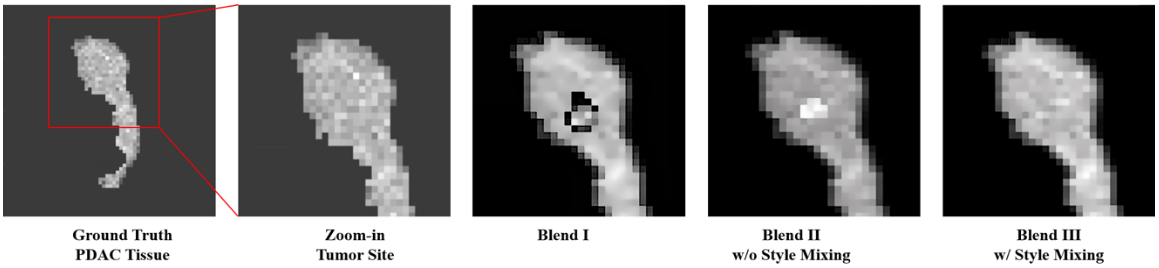


Figure 4.4 Blending Methods Comparison. The comparison of blending methods. The two left images show the details of the tumor site texture in the ground truth data as a visual reference. The rest of the images show the blend of a tumor into the healthy pancreas tissue with different blending methods. We can observe that Blend III has the best visual similarity.

Table 4.3 Slice-wise FID Values of Blending Methods

Blending Methods	FID-Sag	FID-Ax	FID-Cor
Blend I	42.10	40.26	32.94
Blend II	21.01	35.82	12.62
Blend III	13.21	13.88	10.06

3.3. Enhanced training dataset with synthesized data to improve 3D PDAC tumor classification

We trained a 3D CNN classifier using two different dataset configurations, with or without adding the synthetic data and compared the performance of the binary classification of PDAC between them. Adding synthesized data to the training data enlarged the training data. It reduced the imbalance between positive and negative categories because, in practice, it is usually more challenging to access PDAC patient images than healthy pancreas images. We had 174 PDAC tumor images and 254 healthy pancreas tissue images (a combination of 80 TCIA pancreas CT data and a non-tumorous portion of 174 PDAC data), all from real-world CT scans. All the input images were 1 mm isotropic resolution CT volume and trimmed to a size of 64×64×64. Of all the data, 35 tumorous pancreas images and 51 healthy pancreas images, i.e., 20% of all data, were saved as the test dataset. We had two configurations for the training dataset. The first configuration (Config I) only contained real data for training, i.e., 139 tumorous pancreas images and 203 healthy pancreas images. The second configuration (Config II) included the training set from Config I and synthesized data, i.e., 114 synthetic PDAC images and

50 synthetic healthy pancreas images. Both training datasets had the same test dataset for comparison. Config I was a baseline for real-world imbalanced data having a smaller size, and Config II used synthetic data to balance the entire dataset. **Table 4.4** summarizes the training used in the two different configurations. For data augmentation, each CT scan underwent a random rotation along a single axis. The rotation angle was randomly selected from a set of options: 5°, 10°, 20°, and 40°. The direction of rotation, either clockwise or counterclockwise, was also randomly determined. The best parameters of each model were found with a grid search of a parameter space consisting of batch size and learning rate, where the batch sizes included 8, 12, and 16, and the learning rates included 0.001, 0.0001, and 0.00001. All the models were trained with an NVIDIA RTX 3090 GPU and validated using three-fold cross-validation.

Table 4.4 Dataset Configurations for Classifier Experiments

	Training Set
Config I	139 True PDAC 203 True Healthy Pancreas
Config II	139 True + 114 synthesized PDAC 203 True + 50 synthesized Healthy Pancreas

Figure 4.5 shows the receiver operating characteristic (ROC) and precision-recall (PR) curves for the classification models. Config I had an ROC AUC (area under the curve) value of 0.67 and a PR AUC value of 0.80, while Config II had an ROC AUC

value of 0.79 and a PR AUC value of 0.87. Given that the model was trained on a limited dataset, we assessed the classifier's accuracy using three-fold cross-validation. In Configuration I, the average accuracy was 0.57 with a standard deviation of 0.07, while in Configuration II, the average accuracy was 0.67 with a standard deviation of 0.13. The analysis of the results indicated that, as the training dataset was enlarged and the training data in the two classes became more balanced, there was an observable increase in the AUC and precision–recall metrics. This finding implies that including synthesized data to solve issues related to training data quantity and class imbalances had a beneficial effect on the performance of the PDAC classifier. It is worth acknowledging that, despite advancements in utilizing large quantities and balanced training data, there is considerable room for enhancing the classifier's overall performance. This might be achieved using a purpose-built 3D CNN model or further refining the training methodology. By leveraging synthesized data, conducting extra research and analysis of the identification of supplementary components could improve the performance of classifiers and yield superior outcomes.

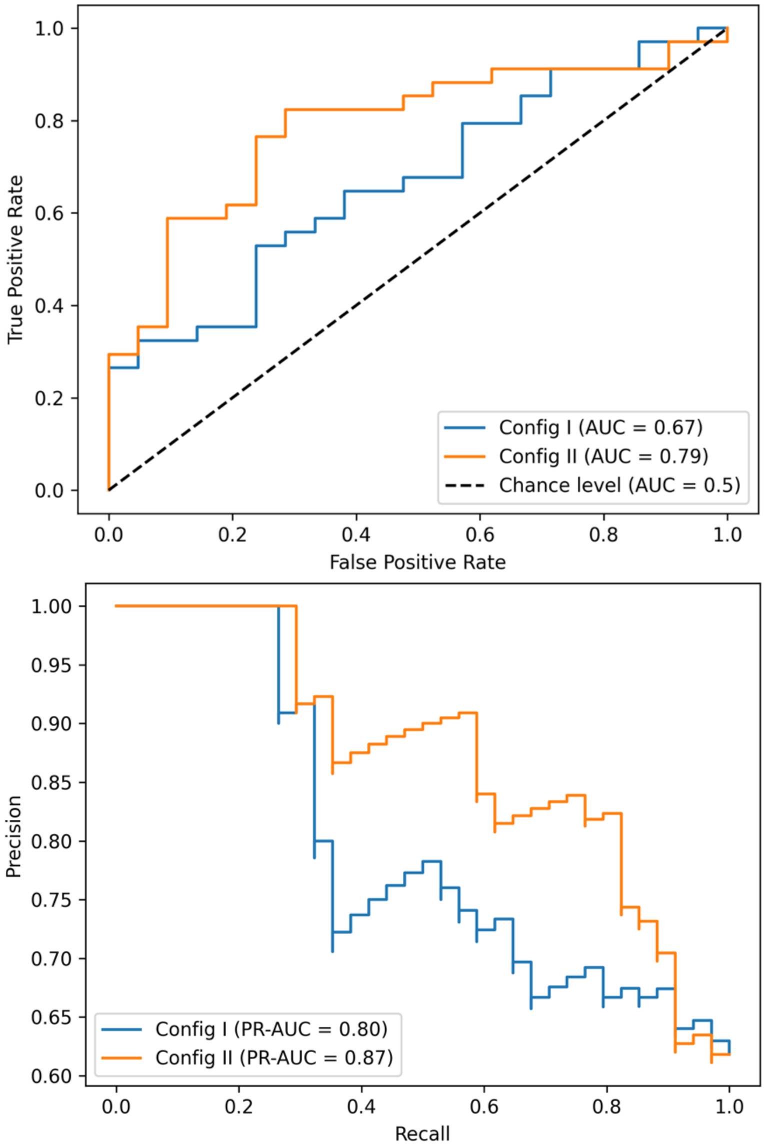


Figure 4.5 3D CNN Classifier Performance. ROC curves (top) and PR curves (bottom) with two configurations of training datasets.

4. Discussion

In this work, we developed a 3D GAN model, 3DGAUnet, for tumor CT image synthesis, compared different blending methods for CT image synthesis, and explored the impact of our synthesis method on a real-world 3D CNN classifier for tumor diagnosis.

3DGAUnet was explicitly designed to synthesize clinical CT images by combining the 3D U-Net architecture with GAN principles to generate realistic 3D CT scans of clinical data. To ensure its accuracy, we trained the model using 3D image data of tumor and healthy pancreas tissue. The quality of the synthesized images was rigorously evaluated using both qualitative and quantitative methods. The generated images demonstrated a more realistic texture than the general 3D-GAN with a CNN-based generator and exhibited the advantage of preserving spatial coherence better than 2D methods. One notable feature of our 3DGAUnet model was its ability to learn the inter-slice gradient, contributing to the overall realism of the generated data. The model also showcased consistent 2D FID values across all three axes, further affirming its capability to produce high-quality 3D images.

3DGAUnet uses preprocessed, fixed-size image cubes. Preprocessing still requires significant human labor and judgment, such as eliminating defects caused by high-density material markers and creating standardized volumes for each training dataset. Medical professionals must manually annotate all training datasets. More automatic methods would be desirable to reduce the significant cost of acquiring data for model training. The model-generated 3D volumes do not come with the metadata a standard DICOM file contains, which means the affine alignment information is still missing. This would limit the potential of downstream applications, such as using the synthesized 3D volumes to enlarge a dataset to improve the training of whole-body segmentation tasks.

The blending method used in this work can insert the tumor image into the background tissue image at a fixed location. For the use case involving mesenchymal organs, the various locations of the tumor have distinct anatomical meanings, and the model must also acquire this information. One possible extension would be to include a segmentation module that can extract the features of each tissue type or organ from the original CT scans. Adopting this approach might eliminate the necessity of considering the blending technique, potentially mitigating the occurrence of faults.

The F3D score we implemented in this work for evaluation is a naive extension of the original FID metric. The activation vectors were extracted from an untrained, cold-started 3D CNN model. The stability of the F3D score in different image domains still needs to be tested. In the future, a benchmark dataset and rigorous testing procedures will be necessary to establish a standardized measure.

The 3D CNN classifier simplifies the clinical diagnosis issue into a binary classification challenge because only healthy normal pancreas images and pancreas images contain a tumor contrast. Multiple conditions, such as non-cancerous lesions, inflammatory or metastatic lesions, and vascular abnormalities, may coexist with the pancreas. Due to a lack of high-quality, well-annotated data, these obstacles still need to be considered and will increase the cost of building such a model.

5. Conclusion

The 3DGAUnet model represents a significant advancement in synthesizing clinical tumor CT images, providing realistic and spatially coherent 3D data, and it holds tremendous potential for improving medical image analysis and diagnosis. In the future, we will continue to address problems associated with computer vision and cancer. We want to investigate reasonable usages of synthetic data and evaluations of data quality and usability in practice, for example, their effectiveness in training reliable classification models. We also plan to conduct reader studies involving domain experts (e.g., radiologists) to assess whether synthetic CTs can enhance diagnostic accuracy. This would offer additional clinical validation regarding the resemblance of synthesized CTs to real-world PDAC tumor characteristics.

From the projects in Chapter 3 and this chapter, I have successfully achieved the second objective of this dissertation by creating an advanced model that generates high-quality 3D cancer image data on PDAC. This development significantly bridges the gap in current research, providing valuable tools for furthering our understanding and improving diagnostic AI technique development.

6. References

[1] American Cancer Society. Cancer Facts & Figures 2022. American Cancer Society; Atlanta, GA, USA: 2022. Technical Report.

[2] Program S.R. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Populations-Total U.S. (1969–2020) [Katrina/Rita Adjustment]-Linked To County Attributes-Total U.S., 1969–2020 Counties. National Cancer Institute, DCCPS; Bethesda, ML, USA: 2022. [(accessed on 1 January 2023)]. Technical Report. Available online: www.seer.cancer.gov

[3] Wu J., Qian T. A survey of pulmonary nodule detection, segmentation, and classification in computed tomography with deep learning techniques. J. Med. Artif. Intell. 2019;2 doi: 10.21037/jmai.2019.04.01.

[4] Tandel G.S., Biswas M., Kakde O.G., Tiwari A., Suri H.S., Turk M., Laird J.R., Asare C.K., Ankrah A.A., Khanna N., et al. A review on a deep learning perspective in brain cancer classification. Cancers. 2019;11:111. doi: 10.3390/cancers11010111.

[5] Sharif M.I., Li J.P., Naz J., Rashid I. A comprehensive review on multi-organs tumor detection based on machine learning. Pattern Recognit. Lett. 2020;131:30–37. doi: 10.1016/j.patrec.2019.12.006.

[6] Radiya K., Joakimsen H.L., Mikalsen K.Ø., Aahlin E.K., Lindsetmo R.O., Mortensen K.E. Performance and clinical applicability of machine learning in liver computed

tomography imaging: A systematic review. Eur. Radiol. 2023;33:6689–6717. doi: 10.1007/s00330-023-09609-w.

[7] Xue Y., Tong W., Neri F., Zhang Y. PEGANs: Phased Evolutionary Generative Adversarial Networks with Self-Attention Module. Mathematics. 2022;10:2792. doi: 10.3390/math10152792.

[8] Baltruschat I.M., Nickisch H., Grass M., Knopp T., Saalbach A. Comparison of deep learning approaches for multi-label chest X-ray classification. Sci. Rep. 2019;9:6381. doi: 10.1038/s41598-019-42294-8.

[9] Coudray N., Ocampo P.S., Sakellaropoulos T., Narula N., Snuderl M., Fenyö D., Moreira A.L., Razavian N., Tsirigos A. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nat. Med. 2018;24:1559–1567. doi: 10.1038/s41591-018-0177-5.

[10] Li Y., Zhang H., Xue X., Jiang Y., Shen Q. Deep learning for remote sensing image classification: A survey. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 2018;8:e1264. doi: 10.1002/widm.1264.

[11] Razzak M.I., Naz S., Zaib A. Classification in BioApps: Automation of Decision Making. Springer; Cham, Switzerland: 2018. Deep learning for medical image processing: Overview, challenges and the future; pp. 323–350.

[12] Li S., Song W., Fang L., Chen Y., Ghamisi P., Benediktsson J.A. Deep learning for hyperspectral image classification: An overview. IEEE Trans. Geosci. Remote Sens. 2019;57:6690–6709. doi: 10.1109/TGRS.2019.2907932.

[13] Chu L.C., Park S., Kawamoto S., Yuille A.L., Hruban R.H., Fishman E.K. Pancreatic cancer imaging: A new look at an old problem. Curr. Probl. Diagn. Radiol. 2021;50:540–550. doi: 10.1067/j.cpradiol.2020.08.002.

[14] Si K., Xue Y., Yu X., Zhu X., Li Q., Gong W., Liang T., Duan S. Fully end-to-end deep-learning-based diagnosis of pancreatic tumors. Theranostics. 2021;11:1982. doi: 10.7150/thno.52508.

[15] Foret P., Kleiner A., Mobahi H., Neyshabur B. Sharpness-aware minimization for efficiently improving generalization. arXiv. 20202010.01412

[16] Wei Z., Chen Y., Guan Q., Hu H., Zhou Q., Li Z., Xu X., Frangi A., Chen F. Pancreatic Image Augmentation Based on Local Region Texture Synthesis for Tumor Segmentation; Proceedings of the 31st International Conference on Artificial Neural Networks; Bristol, UK. 6–9 September 2022; Cham, Switzerland: Springer; 2022. pp. 419–431.

[17] Guan Q., Chen Y., Wei Z., Heidari A.A., Hu H., Yang X.H., Zheng J., Zhou Q., Chen H., Chen F. Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN. Comput. Biol. Med. 2022;145:105444. doi: 10.1016/j.compbiomed.2022.105444.

[18] Wu J., Zhang C., Xue T., Freeman B., Tenenbaum J. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling; Proceedings of the Advances in Neural Information Processing Systems; Barcelona, Spain. 5–10 December 2016.

[19] Antonelli M., Reinke A., Bakas S., Farahani K., Kopp-Schneider A., Landman B.A., Litjens G., Menze B., Ronneberger O., Summers R.M., et al. The medical

segmentation decathlon. Nat. Commun. 2022;13:4128. doi: 10.1038/s41467-022-30695-9.

[20] Roth H., Farag A., Turkbey E.B., Lu L., Liu J., Summers R.M. The Cancer Imaging Archive. NCI; Bethesda, MD, USA: 2016. Data From Pancreas-CT (Version 2)

[21] Hounsfield G.N. Computed medical imaging. Science. 1980;210:22–28. doi: 10.1126/science.6997993.

[22] Oktay O., Schlemper J., Folgoc L.L., Lee M., Heinrich M., Misawa K., Mori K., McDonagh S., Hammerla N.Y., Kainz B., et al. Attention u-net: Learning where to look for the pancreas. arXiv. 20181804.03999

[23] Zhang L., Wen T., Shi J. Deep Image Blending; Proceedings of the IEEE Winter Conference on Applications of Computer Vision; Snowmass Village, CO, USA. 1–5 March 2020; pp. 231–240.

[24] Kikinis R., Pieper S.D., Vosburgh K.G. Intraoperative Imaging and Image-Guided Therapy. Springer; Berlin/Heidelberg, Germany: 2013. 3D Slicer: A platform for subject-specific image analysis, visualization, and clinical support; pp. 277–289.

[25] Heusel M., Ramsauer H., Unterthiner T., Nessler B., Hochreiter S. Gans trained by a two time-scale update rule converge to a local Nash equilibrium; Proceedings of the Advances in Neural Information Processing Systems; Long Beach, CA, USA. 4–9 December 2017;

[26] Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computer vision; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; Las Vegas, NV, USA. 26 June–1 July 2016; pp. 2818–2826.

[27] Gretton A., Borgwardt K.M., Rasch M.J., Schölkopf B., Smola A. A kernel two-sample test. J. Mach. Learn. Res. 2012;13:723–773.

[28] Odena A., Olah C., Shlens J. Conditional image synthesis with auxiliary classifier gans; Proceedings of the International Conference on Machine Learning; PMLR, Sydney, Australia. 6–11 August 2017; pp. 2642–2651.

[29] Tran D., Bourdev L., Fergus R., Torresani L., Paluri M. Learning spatiotemporal features with 3d convolutional networks; Proceedings of the IEEE International Conference on Computer Vision; Santiago, Chile. 7–13 December 2015; pp. 4489–4497.

CHAPTER 5. EFFICIENT VISION TRANSFORMER-BASED FEATURE EXTRACTION MODEL WITH CONTOUR-GUIDED ATTENTION TO IMPROVE CLASSIFICATION MODEL ACCURACY

1. Introduction

Conventional radiology is ineffective in treatment response evaluation for pancreatic cancer; images and biopsies, however, can be used together as definitive diagnostic tools [1]. CT has been the most common cancer diagnostic imaging approach. Conventional radiology for pancreatic cancer has limited capability in early diagnosis and prognosis because images are typically used as a visual reference and provide limited insight into the pattern, texture, and intensity. Radiomics, which extracts quantitative features from the image, significantly increases the throughput of the quantifiable representation of each image screening, ensuring better differentiation with statistical analysis [2].

With the advancement of omics studies and the availability of large data storage and GPU computing, deep learning-based approaches have become feasible in diagnostic applications. Deep learning, as a subdiscipline within the field of AI, has now been considered a state-of-the-art approach for solving 2D computer vision (CV) tasks in medical imaging, such as segmentation, registration, or classification of chest X-rays and tissue histopathology images [3][4]. Images generated by CT, MRI, and their variant technologies are currently the dominant modality of medical imaging data. Both CT and MRI images are volumetric data, a 3D representation of the region of interest. The way the object is distributed in the 3D space is modeled by uniformly sized primitives called voxels. The advantage of using volumetric information in CT or MRI data is that it preserves the context from adjacent slices that 2D models cannot learn efficiently. A better approach is needed for handling 3D imaging data.

Vision transformer (ViT) is an architecture that effectively solves 2D CV tasks, with a trade-off on high computational demand due to its attention mechanism. Transformer attention architecture was initially proposed for language translation tasks by assigning weights to the vector representation of words. Meanwhile, a positional encoding shows the word's position in the sequence [5]. The attention mechanism estimates the relative importance of the word. The language translation task will compare the keyword and the query word to discriminate if both words are equivalent in relative significance in a sequence. Vision transformer will use the image patch as input, as the word in a sentence, and position embedding is the coordinates of the patch. The original multi-head attention mechanism in ViT exhaustively compares neighboring pixels in the whole image. The smaller the size of the attention patch, the higher the resolution and more computational overhead.

In a cancer CT image, when tumor volume was extracted from labeled contour as input data, neglecting the padded empty volume is favorable since the pixels in the padding are not related to the physical existence. Hence, sparse attention is favorable to identify key sampling points that reduce the training cost, leading to theoretical convergence. Recent studies in CV have reported promising approaches using either predefined local attention [6] to improve the performance or reduce the computation complexity in ViT applications. Still, none of them is suitable for CT images. This study designed and embedded a novel sparse attention algorithm tailored to the CT image data in our transformer architecture.

Machine learning models with radiomic features are typically used to predict pancreatic cancer outcomes since explainability is better than the black box of deep learning methods. The method introduced in this study utilizes deep hybrid learning, incorporating clinical demographical information, radiomic features, and image features extracted from transformer attention heads. All features are integrated into a GBM. GBM produces a prediction model from an ensemble of weak prediction models, typically decision trees. It builds the model in a stage-wise fashion and generalizes it by allowing optimization of an arbitrary differentiable loss function [7]. While it is flexible as it can be used with different loss functions and base models, it is relatively robust to overfitting, especially if the base models are decision trees. It can handle missing values and categorical variables. Most importantly, GBMs are more interpretable than neural networks. They are easier to understand how each weak model contributes to the final prediction. All these characteristics make them a favored practice in the clinical use case.

In this chapter, I expanded upon my methodology for cancer imaging feature extraction by employing contour-guided attention to accommodate the properties of medical imaging modalities like CT or MRI. Furthermore, for a test, I used the novel features integrate with the radiomic features and clinical features to improve the classifier on the PDAC survival prediction, to address my first objective of this dissertation.

2. Materials and Methods

2.1. Data collection and preprocessing

Patient demographic, treatment, and outcome data were collected from the medical records of patients with PDAC who received stereotactic body radiotherapy (SBRT) between 2007 and 2016 under the UNMC rapid autopsy program. Seventy-three of the patients have clinical information paired with contrast-enhanced CT images. All these patients used the same protocol to receive the contrast-enhanced abdomen CT scan. The clinicians manually reviewed and adjusted each patient's tumor contour, following the same criteria to maintain consistency. The information includes the patient's date of birth, date of diagnosis, gender, tumor location, Tumor-Nodes-Metastsis (TNM) stage at diagnosis, date of radiation delivery, and any chemotherapy or surgery delivered concurrently with the radiation treatment. All data collection was approved by the Institutional Review Board (IRB) of our institution (Protocols: 728-16-EP and 127-18-EP). For the stereotactic body radiotherapy (SBRT), all patients received 25–40 Gy in five fractions with fiducial markers to localize and track the tumor. Overall survival was calculated from the time of the SBRT simulation CT scan to the date of death. The disease recurrences were registered with two consecutive follow-up CT scans and MRI scans after the delivery of radiation therapy.

Among the 73 available CT image stacks, while the x-axis and y-axis are at 512 pixels resolution, the slice thickness varies. Hence, all the CT images were resampled to the 2mm thickness of the slice, ensuring the volumes extracted from the CT image stacks were under the same scale on each axis. Each voxel is 2mm cubed with Slicer (Ver.

5.0.2). The 3D volume of the tumor in the NIfTI format was created with SimpleITK (Ver. 2.1) and NiBabel (Ver.4.0). Each tumor volume has a dimension of 128 pixels cubed with zero padding (**Figure 5.1**).

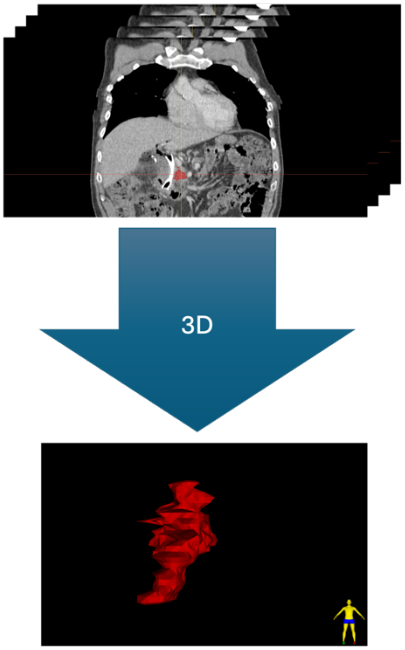


Figure 5.1 3D Volume Extraction from Original DICOM Files.

2.2. Contour-guided attention image feature extraction

In the original ViT architecture, the input image is processed into non-overlap patches, each with a pixel dimension of 16×16. The size of each patch came from the decision to conserve the memory and computing power and the granularity that improves the performance. For instance, if the original input image has a pixel dimension of 512×512, with the ViT, 1024 non-overlap patches will be generated, then flattened and linearly embedded with the location index. Each patch is treated as a token for the transformer. The transformer's complexity comes from its self-attention mechanism, which computes interactions between all pairs of tokens. The computational complexity is $O(n^2 \cdot d)$, where n is the number of patches and d is the dimension of each token embedding. Therefore, increasing the number of patches increases the square of the computation required in self-attention layers, leading to higher computation overhead. For the performance of CV tasks, the bigger patch would limit the transformer recognizing fine details within each patch. The original ViT was introduced for natural image tasks; 16×16 patches are good enough to separate everyday objects (e.g., cars, humans, etc.). In the application of medical imaging, the diagnostic decision is made based on pixel-level observation from the clinician; the patch size needs to be changed to smaller, such as 4×4 or 1×1. In this case, the input image with a dimension of 512×512 would generate 16,384 4×4 patches or 262,144 1×1 patches, making the computation too expensive. Therefore, a better approach should be taken to resolve the issue.

In the original ViT, the attention mechanism has a high capacity by querying every token of the entire image (the size of the d mentioned above), which has a high computation cost and is hard to converge. It requires longer training epochs and substantial data augmentation. However, since large medical image datasets are still barely available and expensive to generate, the advantage of using transformer attention architectures is reduced. ViT variants, such as Swin Transformer [8] or Deformable Attention (DAT) [9], use shifted windows and reference points to achieve sparse attention. These two approaches inspired my attention mechanism.

For a natural image (such as a picture of a dog sitting on the lawn), the depth of field is warped into a 2D projection. On the contrary, for a medical image (e.g., from CT or MRI scan), since every object in the 2D slide is on the same focal plane, a binary mask can easily find the object's contour. In our use case, tumor volume or any region of interest is extracted from the original scan and zero-padded. For each 2D slide, a binary contour mask can separate the zero-padded and tumor tissue areas. Patches containing only zero padding in the image will be dropped from the attention field (**Figure 5.2**). It reduces the sizes of n and d, and decreases the computational complexity. In DAT, deformable points that interpolate from the reference points and offset value could lead to explainability loss. My approach circumvents this issue, and is more appropriate for the medical imaging applications.

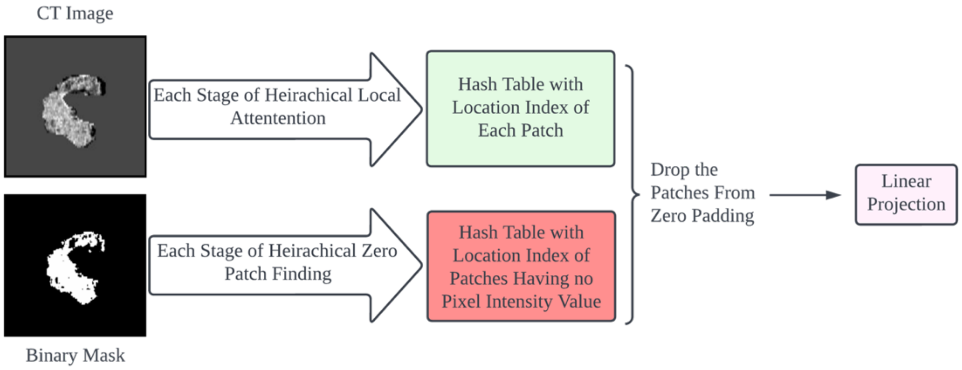


Figure 5.2 Schematic View of Contour-guided Attention.

As shown in **Figure 5.3**, the input image with a pixel dimension of 128×128 is split into 16×16 patches with position embeddings, and the patches containing tissue images within the contour are kept for the attention module. In the second stage, the input image is divided into 8×8 patches; only the patches containing the tissue images are kept. The last stage will have the patches with a dimension of 4×4. Then, a linear projection layer will transform all the kept patches into tokens. A multi-layer perception (MLP) network will take all the tokens and train on the regression loss of survival risk. When the loss converges, tokens with a top attention score will be selected as the output of this feature extraction method. The total number of image patches from one input image is 1344 (**Table 5.1**).

Table 5.1 Image Token Dimensions in Each Stage

	Patcth Size	Feature Dim.	No. of Features
Input Image	-	-	-
Stage I	16 × 16	256	64
Stage II	8 × 8	64	256
Stage III	4 × 4	16	1024

With contour-guided attention, it could be reduced to less than 750 tokens. Since the patches are acquired with a multi-scale scheme, it can extract features from larger areas and fine details without significantly increasing the computational demand.

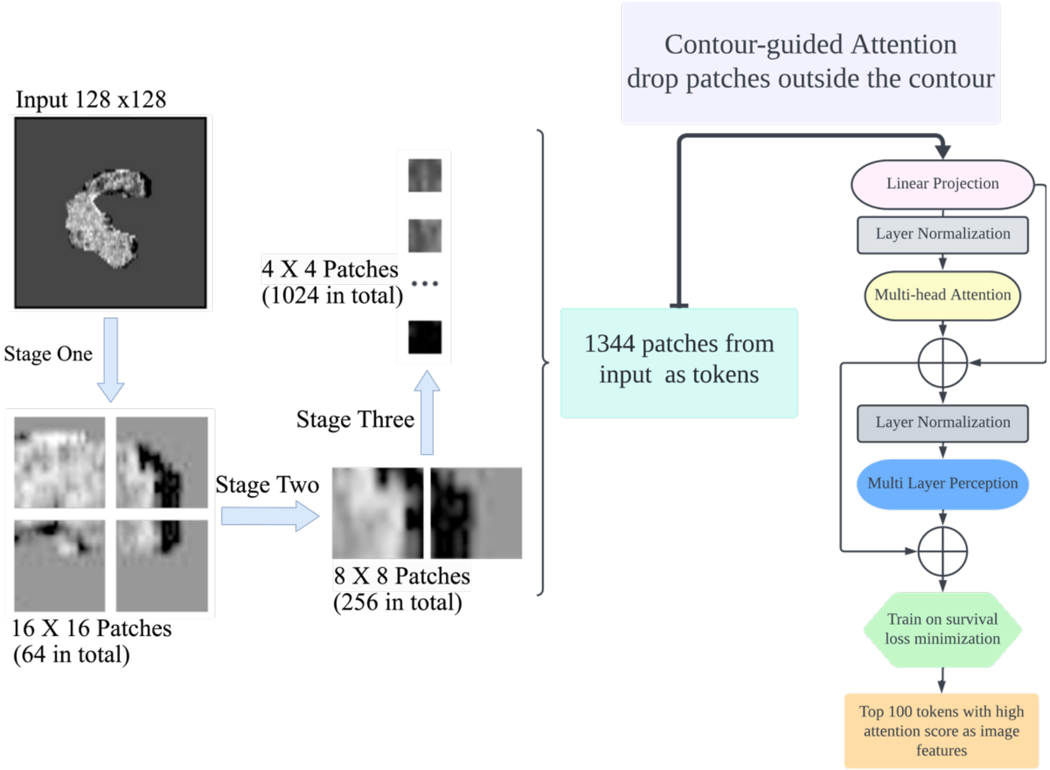


Figure 5.3 ViT Feature Extractor with Contour-guided Attention.

2.3. Hybrid model with GBM

A GBM model is used to integrate multi-modal features and is designed to predict the survival risks. The feature set includes the top 100 image features with the highest attention scores from the ViT-based model, 170 radiomic features calculated from each CT image with low covariance, and demographic features collected from clinical records, including age, sex, tumor position, T-stage, and concurrent chemotherapy (**Table 5.2**).

Table 5.2 Features used for GBM

Feature Source	Data Type
Top 100 Attention Tokens	Vector
170 Radiomic Features	Numerical
Gender	Catagorical
Age	Numerical
T Staging	Catagorical
N Staging	Catagorical
Chemotherapy	Catagorical

The GBM uses an objective function to estimate the survival risk in survival prediction. Since we must handle right-censored survival data, the Cox proportional hazard (CPH) model is used [10]. The partial likelihood is the product of the probability at each event time that the event has occurred to an individual i:

$$\prod_{i:E_i=1} \frac{\exp(\widehat{h_\beta}(x_i))}{\sum_{j\in\Re(T_i)} \exp(\widehat{h_\beta}(x_j))} \tag{12}$$

where T_i is the event time, E_i is the event indicator with '0' for censored patients and '1' for patients at risk, and x_i is the baseline data for i^{th} observation. The product is defined over patients with an observable event E_i=1. The risk set $\Re(t)= \{i: T_i \geq t\}$ is the set of patients still at risk of failure at time t, and $\widehat{h_\beta}(x_i)$ is the linear function estimate the log-risk.

To train a GBM for the Cox proportional hazards model, the objective function used is typically the negative partial log-likelihood of the Cox model. The partial log-likelihood is defined as:

$$l(\beta) = \sum_{i=1}^{n} \delta_i\left(\beta^T x_i - \log \sum_{j=1}^{n} \exp\left(\beta^T x_j\right)\right) \tag{13}$$

where β is the vector of coefficients, x_i is the vector of predictor variables for the i^{th} sample, and $\delta_i$is the event indicator for the i^{th} sample (1 if the event occurred, and 0 otherwise).

The objective function is then minimized during the training to find the optimal values for the coefficients β. The objective function's gradient concerning the coefficients is calculated using backpropagation and used to update the coefficients using gradient descent.

2.4. 3D CNN model

One of our objectives in developing ViT-based feature extraction is to improve the performance of PDAC survival prediction and tumor identification. To test our developed feature extraction method, we built a 3D CNN model. This 3D CNN is a baseline that only uses images without the ViT features, radiomic features, and clinical features to compare with the GBM hybrid model in survival prediction CI and AUC.

We modified the 17-layer 3D CNN model [11] used in the last chapter to suit survival prediction needs by swapping the output layer from classification to a Cox regression layer, with the same loss as the GBM model. In this manner, we would be able to get both 1-year survival AUC and CI from the same CNN architecture.

The 3D CNN classifier has four 3D convoluted blocks (Conv 3D), with the first block consisting of 64 filters followed by 128, 256, and 512 filters, all with a kernel size of 3×3×3. Each Conv3D layer is followed by a max-pooling (MaxPool) layer with a stride of 2, ReLU activation, and batch normalization layer (Batch Norm). This 3D CNN model has four Con3D-MaxPool-BatchNorm blocks and is intended to capture visual features from coarse to fine. The final output first flattens the production of the last convolutional block and passes it to a fully dense layer with 512 neurons. A dropout layer with a tunable dropout rate follows to prevent overfitting.

2.5. Experimental design

The 73-patient dataset was divided into a training set including 51-patient data (70%) and a testing set including 22-patient data (30%). The testing set was withheld for the training/boosting. The contour-guided attention ViT feature extractor was configured to train on a maximum of 50 epochs with early stopping; the optimal batch size was tested on 5, 10, and 20, and the optimal learning rate was evaluated from 0.00001 to 0.01. For the GBM, the optimal learning rate was tested from 0.0001 to 0.1; the optimal number of trees was tested on 20, 50, 100, and 500; the optimal tree depth was tested on 3, 5, 7, and 9. Both training tasks were done with 500 rounds of 3-fold cross-validation. Data augmentation was done with random rotations in each training round.

Since the PDAC has a low 5-year survival rate, in our dataset, no patient has a survival time longer than 3 years. One way to create the prediction model is to make it a binary classification task, with patients having over 1-year survival labeled ‘positive’ and less than 1-year as ‘negative’.

The binary classification performance was calculated from the confusion matrix. Given that TP, TN, FP, and FN correspond to a true positive, true negative, false positive, and false negative, respectively, the results were measured using precision $TP/(TP+FP)$, recall $TP/(TP+FN)$, true positive rates $TPR=TP/(TP+FN)$, and the false positive rates $FPR=FP/(FP+TN)$. The area under the curve (AUC) was calculated from the receiver operating characteristic (ROC) curve, which was plotted as true positive rates against the false positive rates under different cutoffs or as the precision against the recall.

The AUC on 1-year survival and concordance index (CI) were reported as the performance metrics. The CI measures the predictive accuracy of a survival model [12]. It is defined as the probability that the predicted survival probabilities of a random pair of subjects with different survival times are correctly ranked and calculated as follows:

$$\mathrm{C} = \frac{\mathrm{H+S}}{\mathrm{H+D+S}} \tag{14}$$

H is the number of pairs of subjects where the predicted survival probabilities are correctly ranked. D is the number of pairs where the expected survival probabilities are not ranked correctly. S is the number of pairs where the survival times are the same.

A CI of 1 indicates perfect predictive accuracy, while a value of 0.5 indicates that the predictions are no better than random. Values of CI above 0.5 indicate that the model is making valuable predictions, while values below 0.5 indicate that the model is not performing well. CI and AUC are both metrics used to evaluate the performance of predictive models, but CI is a measure of the discriminatory power of a risk prediction model. It is commonly used in survival analysis to evaluate how well the model can discriminate between different outcomes over time. Often used in survival analysis to evaluate prognostic models, such as Cox proportional hazards models.

3. Results

3.1. Contour-guided attention ViT feature extraction

With contour-guided attention, image tokens (features) were extracted in each stage (**Figure 5.3**), linear projected, and concatenated. The tokens with the top 100 attention scores were kept as features for the GBM classifier. One benefit of the ViT feature is that the top-scored attention can be mapped back to the original CT images. This provides the model explainability and demonstrates the relationship between model decision and anatomy.

Figure 5.4 visualizes the tokens with high attention scores with a heatmap. The yellow color highlights that location is important to the model when making survival predictions.

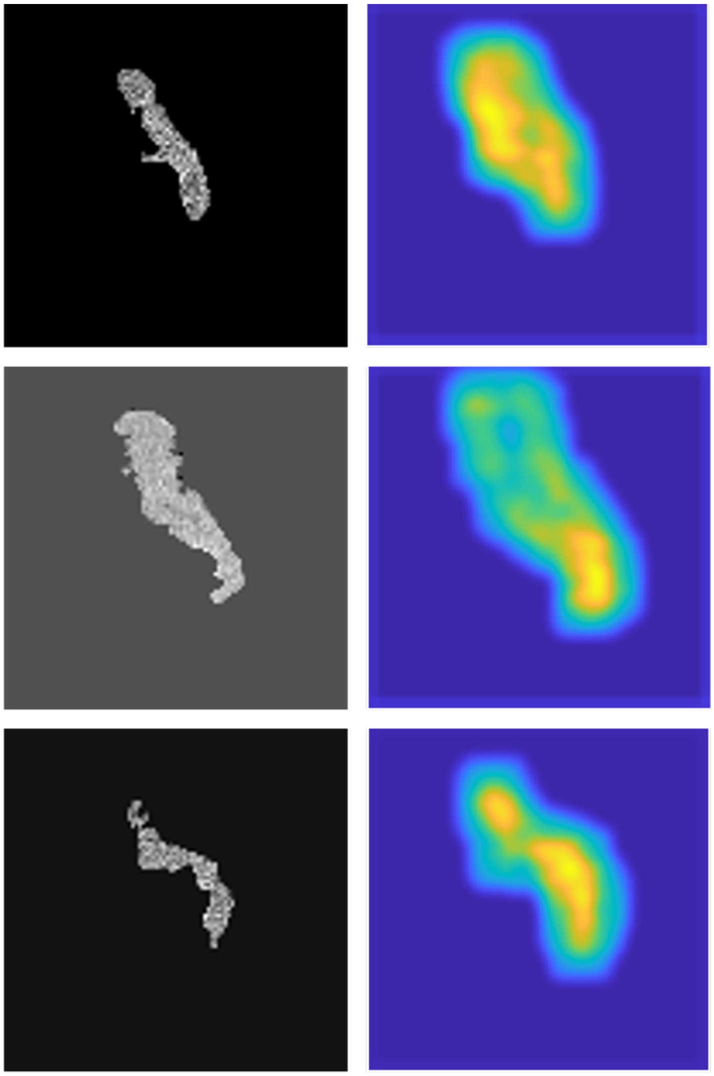


Figure 5.4 Heatmap of Attentions for Three Patients. The top 100 tokens with high attention scores are visualized in the Heatmap

3.2. GBM model on survival prediction

In total, 275 features were fed into the GBM model. The CI values were used to compare the prediction performance of the GBM model against the 3D CNN model that only used the CT volumetric data as predictors. As shown in **Table 5.3**, the GBM model has better performance than baseline 3D CNN model. From the comparison of the AUC on 1-year survival prediction (**Figure 5.5**), the GBM model with multi-modal features showed a consistent performance advantage over the 3D CNN model.

Table 5.3 C-Index Comparison

Model	C-Index
GBM	0.673
3D CNN	0.592

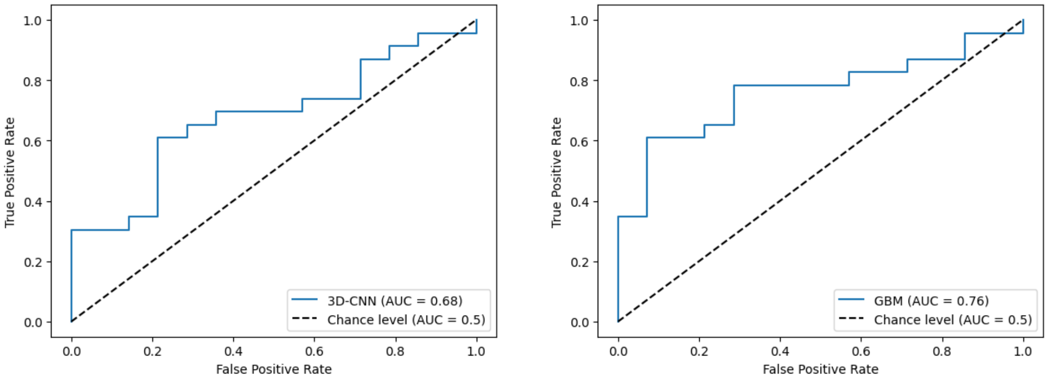


Figure 5.5 AUC on 1-year Survival Predictions.

4. Discussion

This study explores the development of a ViT-based feature extraction module combined with radiomic features and clinical information to predict survival in PDAC. PDAC presents significant diagnostic challenges due to its asymptomatic early stages and high fatality rates. Conventional imaging techniques often fail to detect early-stage cancer, highlighting the need for advanced methodologies. This study introduced a first-ever contour-guided attention mechanism to the ViT architecture to tailor it for CT image data. The model reduced computational complexity and improved learning efficiency by focusing “attention” on relevant image patches (tumor regions) and ignoring noninformative areas (padding). This approach mimics the clinical review process, where attention is directed toward specific anatomical structures. By integrating radiomic features and clinical demographic features, this hybrid model showed promising accuracy results with our small dataset. However, it needs to be further tested for its generalizability with larger datasets, although such datasets are not yet available.

Another aspect worth looking into is how to use the attention score on image tokens better and enhance model explainability since the visualized attention is supposed to connect the model decision and the anatomy. From the attention map presented in the result, the heatmap seems to display a greater area than the actual contour with high attention scores. When it is close to the edge of the contour, how to interpret the connection between morphology and model decision is still questionable. A more sophisticated visualization technique could be designed incorporating advice from experienced clinicians.

5. Conclusion

Key findings from this study are as follows. The contour-guided attention mechanism efficiently extracted features from ViT architecture without increasing computational overhead. The hybrid model integrating ViT-extracted features, radiomic features, and clinical features provides a robust framework for predicting PDAC outcomes despite the limited dataset size, emphasizing its potential for clinical application.

To guarantee the model's generalizability and investigate its value in the visualization of attention maps, future research should validate the model on larger datasets. This study lays the groundwork for advanced, reliable early detection methods in pancreatic cancer, potentially improving patient outcomes through timely intervention. This work is related to the Objective One of this dissertation.

6. References

[1] Tummala, P., Junaidi, O., & Agarwal, B. (2011). Imaging of pancreatic cancer: An overview. Journal of Gastrointestinal Oncology, 2(3), 168.
[2] Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., De Jong, E. E., Van Timmeren, J., ... & Walsh, S. (2017). Radiomics: the bridge between medical imaging and personalized medicine. Nature reviews Clinical oncology, 14(12), 749-762.
[3] Baltruschat, I. M., Nickisch, H., Grass, M., Knopp, T., & Saalbach, A. (2019). Comparison of deep learning approaches for multi-label chest X-ray classification. Scientific reports, 9(1), 6381.
[4] Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., ... & Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nature medicine, 24(10), 1559-1567.
[5] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.
[6] Arar, M., Shamir, A., & Bermano, A. H. (2022). Learned queries for efficient local attention. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10841-10852).
[7] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of Statistics, 1189-1232.
[8] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: A hierarchical vision transformer uses shifted windows. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 10012-10022).
[9] Xia, Z., Pan, X., Song, S., Li, L. E., & Huang, G. (2022). Vision transformer with deformable attention. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 4794-4803).
[10] Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187-202.
[11] Tran D., Bourdev L., Fergus R., Torresani L., Paluri M. Learning spatiotemporal features with 3d convolutional networks; Proceedings of the IEEE International Conference on Computer Vision; Santiago, Chile. 7–13 December 2015; pp. 4489–4497.
[12] Chen, W., Hou, X., Hu, Y., Huang, G., Ye, X., & Nie, S. (2021). A deep learning and CT image-based predictive model for the prediction of survival in non-small cell lung cancer. Medical Physics, 48(12), 7946-7958.

CHAPTER 6. CONCLUSION, IMPACTS, AND FUTURE DIRECTIONS

In this dissertation, significant progress was made in bridging the gaps identified in the research by conducting comprehensive studies and integrating clinical data. The associations between quantitative image features and clinical events were elucidated for the first time with the study on prostate cancer and exposure to Omega-3 and statin (Chapter 2). The synthesis of high-quality 2D medical imaging and cancer volumetric data was advanced by implementing cutting-edge generative models (Chapters 3 and 4). Furthermore, a new variety of image features were extracted by leveraging the Vision Transformer (ViT) architecture with noval contour-guided attention (Chapter 5). The integration of multi-modal data, including radiomic features, ViT features, and clinical information, resulted in improved model accuracy, demonstrating the benefits of a holistic approach to medical image analysis.

Through these contributions, my research aimed to significantly enhance the tools available for cancer imaging feature extraction and synthesis. This potent improvement can enhance cancer diagnosis, prognosis predictions, and overall radiotherapy image analysis. Ultimately, it has the potential to enhance patient outcomes and advance the field of medical imaging analysis.

Despite these achievements, there are still challenges and limitations to be addressed. For instance, while the models developed have shown promising results, their scalability to larger datasets and generalization to different types of medical imaging data need further validation. Other architectures could potentially enhance the quality of synthesized cancer images. Furthermore, the potentials and limitations of prediction models with multi-modal features need more rigorous investigation.

In forthcoming research, I intend to keep improving the ViT feature extraction method and the integration of features from other modalities. Enhancement of the explainbility with the multi-modal data will be explored. This endeavor will facilitate the identification of oncogenic genes and the detection of biomarkers, significantly advancing our understanding of cancer pathogenesis.

CHAPTER 7. APPENDIX

Table 7.1 Radiomic Feature List

Feature Class	Feature Name
shape	VoxelVolume
shape	Maximum3DDiameter
shape	MeshVolume
shape	MajorAxisLength
shape	Sphericity
shape	LeastAxisLength
shape	Elongation
shape	SurfaceVolumeRatio
shape	Maximum2DDiameterSlice
shape	Flatness
shape	SurfaceArea
shape	MinorAxisLength
shape	Maximum2DDiameterColumn
shape	Maximum2DDiameterRow
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability

glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis

glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis

gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness

firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity

glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy

glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance

glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis

gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation

firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis

glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm

glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity

glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis

gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared

firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis

glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC

glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis

glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis

glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range

firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity

ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade
glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1

glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean
glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance

glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness
gldm	GrayLevelVariance
gldm	HighGrayLevelEmphasis
gldm	DependenceEntropy
gldm	DependenceNonUniformity
gldm	GrayLevelNonUniformity
gldm	SmallDependenceEmphasis
gldm	SmallDependenceHighGrayLevelEmphasis
gldm	DependenceNonUniformityNormalized
gldm	LargeDependenceEmphasis
gldm	LargeDependenceLowGrayLevelEmphasis
gldm	DependenceVariance
gldm	LargeDependenceHighGrayLevelEmphasis
gldm	SmallDependenceLowGrayLevelEmphasis
gldm	LowGrayLevelEmphasis
glcm	JointAverage
glcm	SumAverage
glcm	JointEntropy
glcm	ClusterShade

glcm	MaximumProbability
glcm	Idmn
glcm	JointEnergy
glcm	Contrast
glcm	DifferenceEntropy
glcm	InverseVariance
glcm	DifferenceVariance
glcm	Idn
glcm	Idm
glcm	Correlation
glcm	Autocorrelation
glcm	SumEntropy
glcm	MCC
glcm	SumSquares
glcm	ClusterProminence
glcm	Imc2
glcm	Imc1
glcm	DifferenceAverage
glcm	Id
glcm	ClusterTendency
firstorder	InterquartileRange
firstorder	Skewness
firstorder	Uniformity
firstorder	Median
firstorder	Energy
firstorder	RobustMeanAbsoluteDeviation
firstorder	MeanAbsoluteDeviation
firstorder	TotalEnergy
firstorder	Maximum
firstorder	RootMeanSquared
firstorder	90Percentile
firstorder	Minimum
firstorder	Entropy
firstorder	Range
firstorder	Variance
firstorder	10Percentile
firstorder	Kurtosis
firstorder	Mean

glrlm	ShortRunLowGrayLevelEmphasis
glrlm	GrayLevelVariance
glrlm	LowGrayLevelRunEmphasis
glrlm	GrayLevelNonUniformityNormalized
glrlm	RunVariance
glrlm	GrayLevelNonUniformity
glrlm	LongRunEmphasis
glrlm	ShortRunHighGrayLevelEmphasis
glrlm	RunLengthNonUniformity
glrlm	ShortRunEmphasis
glrlm	LongRunHighGrayLevelEmphasis
glrlm	RunPercentage
glrlm	LongRunLowGrayLevelEmphasis
glrlm	RunEntropy
glrlm	HighGrayLevelRunEmphasis
glrlm	RunLengthNonUniformityNormalized
glszm	GrayLevelVariance
glszm	ZoneVariance
glszm	GrayLevelNonUniformityNormalized
glszm	SizeZoneNonUniformityNormalized
glszm	SizeZoneNonUniformity
glszm	GrayLevelNonUniformity
glszm	LargeAreaEmphasis
glszm	SmallAreaHighGrayLevelEmphasis
glszm	ZonePercentage
glszm	LargeAreaLowGrayLevelEmphasis
glszm	LargeAreaHighGrayLevelEmphasis
glszm	HighGrayLevelZoneEmphasis
glszm	SmallAreaEmphasis
glszm	LowGrayLevelZoneEmphasis
glszm	ZoneEntropy
glszm	SmallAreaLowGrayLevelEmphasis
ngtdm	Coarseness
ngtdm	Complexity
ngtdm	Strength
ngtdm	Contrast
ngtdm	Busyness

Table 7.2 Prostate/Omega-3 Feature Candidates

Top 10 most used features within 1000 rounds of resampling	Times
wavelet_HLL_glszm_LargeAreaHighGrayLevelEmphasis	143
log_sigma_5_0_mm_3D_firstorder_Mean	127
wavelet_HLH_ngtdm_Busyness	102
wavelet_HLH_glcm_ClusterShade	88
wavelet_HHL_glcm_Imc1	87
log_sigma_5_0_mm_3D_glszm_SmallAreaLowGrayLevelEmphasis	82
wavelet_LLH_glszm_GrayLevelNonUniformityNormalized	74
log_sigma_5_0_mm_3D_glcm_Idmn	73
log_sigma_5_0_mm_3D_glszm_SizeZoneNonUniformityNormalized	68
log_sigma_5_0_mm_3D_glszm_SmallAreaEmphasis	65

Table 7.3 Prostate/Statin Feature Candidates

Top 10 most used features within 1000 rounds of resampling	Times
wavelet_HHH_glcm_Correlation	181
wavelet_HHH_firstorder_Median	170
wavelet_HLL_glcm_Correlation	156
log_sigma_5_0_mm_3D_gldm_SmallDependenceLowGrayLevelEmphasis	135
wavelet_HHL_firstorder_Skewness	133
wavelet_HLH_glcm_Correlation	95
wavelet_LHL_glcm_Idmn	92
log_sigma_5_0_mm_3D_glszm_SmallAreaEmphasis	74
wavelet_LHL_glszm_SmallAreaLowGrayLevelEmphasis	74
wavelet_LLH_firstorder_Skewness	58

Table 7.4 Peripheral/Omega-3 Feature Candidates

Top 10 most used features within 1000 rounds of resampling	Times
wavelet_LLL_firstorder_Kurtosis	191
wavelet_LHH_glszm_SmallAreaLowGrayLevelEmphasis	172
wavelet_HLL_glszm_SmallAreaEmphasis	163
wavelet_HLH_firstorder_Mean	144
wavelet_HLH_firstorder_Median	140
wavelet_LHH_glszm_LowGrayLevelZoneEmphasis	131
log_sigma_5_0_mm_3D_firstorder_Skewness	85
wavelet_HHH_firstorder_Mean	67

wavelet_LLH_glszm_SmallAreaEmphasis	62
wavelet_HHH_glszm_GrayLevelNonUniformityNormalized	58

Table 7.5 Peripheral/Statin Feature Candidates

Top 10 most used features within 1000 rounds of resampling	**Times**
wavelet_HHH_glcm_Correlation	269
wavelet_HHL_firstorder_Median	188
log_sigma_5_0_mm_3D_glszm_SizeZoneNonUniformityNormalized	131
log_sigma_5_0_mm_3D_gldm_DependenceNonUniformityNormalized	128
wavelet_HHL_firstorder_Skewness	113
log_sigma_5_0_mm_3D_glszm_SmallAreaEmphasis	112
log_sigma_5_0_mm_3D_glszm_LargeAreaEmphasis	88
wavelet_HHL_glcm_Idmn	81
wavelet_HHH_firstorder_Median	79
wavelet_HLH_firstorder_Skewness	78

Table 7.6 Top Radiomic Features in ACEi Predictions

	ACEi Top 10 Features	Times
Prostate	wavelet_HLH_firstorder_Skewness	187
	log_sigma_5_0_mm_3D_glszm_LargeAreaLowGrayLevelEmphasis	166
	wavelet_HLL_glcm_Imc1	152
	wavelet_LLH_glszm_SmallAreaLowGrayLevelEmphasis	113
	wavelet_HLL_firstorder_Skewness	108
	wavelet_HLH_glcm_ClusterShade	59
	wavelet_HLH_glszm_SmallAreaEmphasis	44
	log_sigma_5_0_mm_3D_glcm_Imc1	32
	wavelet_HLL_glcm_MCC	22
	log_sigma_5_0_mm_3D_glcm_Imc2	21
Peripheral	wavelet_HHH_gldm_DependenceNonUniformityNormalized	290
	wavelet_HLL_glcm_Imc1	200
	log_sigma_5_0_mm_3D_glszm_SmallAreaLowGrayLevelEmphasis	151
	wavelet_LLH_glszm_SmallAreaEmphasis	130
	wavelet_HLH_firstorder_Skewness	112
	log_sigma_5_0_mm_3D_ngtdm_Coarseness	80
	wavelet_LHH_firstorder_Skewness	42
	log_sigma_5_0_mm_3D_glszm_SmallAreaEmphasis	41
	wavelet_LHL_glcm_Imc2	25
	wavelet_HLL_glcm_MCC	20