Title

Application of Convolutional Neural Networks in Multiparametric MR Imaging to Predict Prostate Cancer Progression

Author

Doan, Duc Huy

Publication Date

2024

Peer reviewed|Thesis/dissertation

Application of Convolutional Neural Networks in Multiparametric MR Imaging to Predict Prostate Cancer Progression

by

Duc Huy Doan

THESIS

Submitted in partial satisfaction of the requirements for degree of

MASTER OF SCIENCE

in

Biomedical Imaging

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Susan Noworolski
070ED2AA0DC3441...

Susan Noworolski

Chair

DocuSigned by:
Matthew Gibbons
[illegible]

Matthew Gibbons

DocuSigned by:
[illegible]

Youngho Seo

DocuSigned by:
[illegible]
7ADB1F8EDFF44B4...

Spencer Behr

Committee Members

ACKNOWLEDGEMENTS

To Dr. Susan Noworolski, this project, and its culmination, would not have been possible without your unwavering support and guidance. Your mentorship has been invaluable, not only in shaping the direction of this research but in my growth as a researcher. You have consistently demonstrated a willingness to invest your time and expertise, offering insightful advice and direction whenever I needed it. Moreover, your understanding and accommodation of my commitments outside of this project is tremendously appreciated, allowing me to maintain a much needed balance. I am profoundly grateful for the role you have played in my academic journey.

To Dr. Matthew Gibbons, Dr. Youngho Seo and Dr. Spencer Behr, I am deeply indebted to each of you for your invaluable contributions to this project. Your willingness to dedicate your time and expertise, providing insightful advice, and guiding me through the intricacies of data analysis and interpretation, has been instrumental in this research. Your help to bridge the gap between theoretical findings and their practical application in clinical settings has enriched the significance of this work. I am deeply appreciative of the guidance and support you have provided me throughout this project.

APPLICATION OF CONVOLUTIONAL NEURAL NETWORKS

IN MULTIPARAMETRIC MR IMAGING

TO PREDICT PROSTATE CANCER PROGRESSION

Duc Huy Doan

ABSTRACT

Prostate cancer progression after radical prostatectomy poses a significant risk to patient health. The ability to predict which patients are at a higher risk of progression is crucial for determining appropriate adjuvant therapies. This study investigates the application of convolutional neural networks (CNNs) to pre-surgical multiparametric MRI (mpMRI) for predicting post-surgical prostate cancer progression. The study utilizes a retrospective patient cohort and explores the performance of different CNN architectures (ResNet and DenseNet), normalization methods, and slice selection techniques. The results demonstrate the potential of CNNs in predicting prostate cancer progression, with the best-performing model achieving an accuracy of 0.712. The study highlights the importance of appropriate image normalization and slice selection methods for optimal performance. The findings suggest that CNNs could serve as a valuable tool for aiding clinical decision-making in prostate cancer management.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADC: Apparent Diffusion Coefficient

CNN: Convolutional Neural Network

DCE: Dynamic Contrast-Enhanced

DWI: Diffusion-Weighted Imaging

MR: Magnetic Resonance

MRI: Magnetic Resonance Imaging

T2W: T2-Weighted

1. INTRODUCTION

Prostate cancer (PCa) is the second most deadly cancer for men, claiming over 35,000 lives annually. Even after radical prostatectomy (RP), there is still a possibility of disease progression, local recurrence or metastases (~20-50% within 10 years after surgery) [1]. Therefore, it is crucial to identify those with aggressive disease early to allow for timely intervention of adjuvant therapies. Nevertheless, these adjuvant therapies are not without toil on the patients in terms of costs, toxicity and discomfort. Therefore, patients should undergo adjuvant therapies only when they really need to. An effective tool is needed to assess the risk of cancer progression to determine this.

A simple blood test, measuring prostate-specific antigen (PSA), is readily available and valuable for screening cancer progression. An increase in PSA after RP is an indicator of biochemical recurrence (BCR) which is the first sign that the cancer might be returning. It suggests that some prostate tissue has already grown. It may be due to benign or cancerous tissue in the prostatic surgical bed or due to local or distant cancer metastases. If due to cancer, identifying this as early is important to initiate treatment.

Biopsy histologic grade is an important indicator used in progression risk assessment but it also has its drawbacks. Even histologically similar prostate tumors can follow significantly different disease courses after primary treatment [3].

Multiparametric MRI (mpMRI) has the potential to overcome these limitations as it is both non-invasive and able to examine the entire prostate. It combines images from

multiple MRI parameters such as: T2-weighted imaging (T2W), diffusion-weighted imaging (DWI) with apparent diffusion coefficient (ADC) maps and dynamic contrast enhanced (DCE) MRI. mpMRI has been widely used for detecting clinically significant cancer and may be able to provide an indicator of which cases are apt to progress post surgery, empowering earlier monitoring and earlier interventions for hisk-risk cases.

In this project, an artificial intelligence technique called convolutional neural networks (CNNs) is applied to learn from the pre-surgical multiparametric MR images to classify patients into high-risk versus low-risk of post-surgical cancer progression. The high-risk patients are candidates for earlier monitoring and for adjuvant therapies.

2. METHODS

2.1 Selecting Patient Data

A retrospective study is performed on a patient cohort who underwent radical prostatectomy. The inclusion criteria is that they must have their mpMRI scan less than 1 year before their surgery and either have cancer progression within 3 years of surgery or a follow-up without cancer progression more than 3 years after the surgery. In total, there are 52 patient cases matching these criteria. Out of these patients, 20 are with progression and 32 are without progression.

2.2 Selecting Convolutional Neural Networks

Convolutional neural networks (CNNs) have emerged as a groundbreaking technology in the field of medical imaging, revolutionizing the way we diagnose and treat various diseases. Inspired by the biological visual cortex, CNNs are a class of deep learning models specifically designed to process and analyze visual data. Their unique architecture allows them to automatically learn hierarchical representations of images, enabling them to detect intricate patterns and features that may not be visible to the human eye [4].

In this project, two CNN models are experimented with and compared: ResNet and DenseNet [5][6]. They are very powerful deep learning models that have shown success in many image classification tasks including those used in medical imaging.

ResNet, short for Residual Network, revolutionized the field of deep learning by introducing the concept of residual connections. ResNet networks are designed with skip connections that allow information flow directly through the network, bypassing some layers. This architectural innovation enabled training of significantly deeper networks than previously possible. ResNet also generally achieves faster convergence and better accuracy than traditional deep CNNs [5].

DenseNet is known for its dense connectivity pattern, where each layer connects to all subsequent layers in a feed-forward fashion. This creates dense blocks where each layer receives feature maps of all preceding layers as input. This strengthens feature propagation and encourages feature reuse, leading to improved gradient flow, and parameter efficiency. DenseNet also allows for the training of very deep networks [6].

2.3 Image Preparation For CNN Input

Step 1 - Standardize Image Dimension And Field Of View

Three MRI image types will be used:

- Coil-corrected T2-weighted (T2W) images
- Diffusion-weighted imaging (DWI) apparent diffusion coefficient (ADC) map
- Dynamic contrast enhanced (DCE) MRI max enhancement slope (ES) map

All the ADC and DCE MRI ES images are processed to match the T2W image's field of view and voxel size. The typical image domain is 512 (180 mm) x 512 (180 mm) x 16

(48 mm). The spatial resolution of the width and height is 0.35mm/voxel and that of the depth is 3mm/voxel.

Prostate images are generally acquired in an oblique axial orientation, but may be ordered from the superior to inferior direction or vice versa. This orientation information is stored in the image header as the dcos3 value. That value refers to the orientation of the gradient vector relative to the superior-inferior (S/I) axis. Our images either have the superior/inferior component of this value very close to 1 (pointing in the superior direction) or very close to -1 (pointing in the inferior direction). To standardize our images, the inferior to superior direction is chosen as the standard direction. For images pointing in the opposite direction, the slice order has to be reversed.

Step 2 - Image Registration To T2W

Registering ADC and DCE MRI images to the T2W image is a crucial step in medical image analysis. This process aligns the different image modalities, ensuring that the corresponding anatomical structures are precisely overlaid. Rigid registration was applied first in-plane, then, in the prostate region, in 3D. This method has shown <2mm error on >95% of cases. Manual correction was applied to remaining cases with larger shifts [7]. This alignment is essential for accurate comparison and integration of information from the different image modalities in the CNN training subsequently.

Step 3 - Image Normalization

The original image data is a 2-byte integer value ranging from 0 to (2^{16}-1). For input into the CNN, these values need to be scaled down to the range from 0 to 1. A regular normalization method would be to set the value of 1 to correspond to the max value, set the value of 0 to correspond to the min value and scale all the values in between. However, that approach may not be optimal. Due to the possible presence of artifacts that result in abnormally bright spots, normalizing that way may make the resulting image abnormally dark.

Therefore, another custom normalization method is proposed. For T2W images, the value of 1 is set to correspond to the 80th percentile value instead of the max value. For ADC and DCE images, since they are absolute values of the map, a different set of min and max values are defined to narrow down the range. As these are 2-byte integer images, the min value is set to 500 and the max value is set to 3000 (corresponding to ADC range of 5x10^{-4} to 3x10^{-3} mm^2/s and DCE range of 50% to 300% of baseline/min). All values below the min value will have a value of 0 and all values above the max value will have a value of 1. All other values in between min and max will be scaled to a value between 0 and 1. In this study, the performance of this custom normalization method is compared with the above regular normalization.

Step 4 - Fit Image Dimension Into CNN

The images are cropped to match the dimension required by the CNN (224 x 224). Horizontally, the 144th to 367th pixels are selected. Vertically, the 159th to 383th pixels are

selected. Both dimensions span 78.75mm. This makes the cropped area just cover the anatomy of the prostate with the immediately surrounding areas. For the 3 channels, data from 3 MRI image types, namely T2W, ADC and DCE, are used for each channel.

As the CNN takes in 2D images, only 1 slice out of many available slices in the patient case is selected. An experiment is conducted to determine what is the best approach to select this slice. On one hand, the middle slice can be automatically picked. On the other hand, the image can be preprocessed to segment where the cancer is and the slice in the middle of the largest cancer volume is manually selected. This is potentially more accurate but also requires a lot more effort in preprocessing.

Step 5 - Image Augmentation

Given that the dataset only has a limited number of images, new images could be artificially generated for learning based on those already present. The method used in this project is to create new images by flipping the original images left and right. The new images are anatomically reasonable because the left and right side of the prostate are symmetrical. Other potential methods include shifting (~5mm), rotating (~2-5^{o}) and stretching/shrinking (~2%) the original images.

2.4 CNN Training And Validation

K-Fold cross validation technique (with K = 4) is used to ascertain the accuracy of our model's prediction. It is a resampling method that uses different portions of the data

to train and validate a model on different iterations. It ensures that prediction performance does not depend on the training and validation set selection.

The data is split into 4 folds and 4 iterations are run. For each iteration:

- One fold is designated as the validation set and the other 3 as the training set.
- So 75% of the data is used for training and 25% for validation.
- The training set has left-right augmentation while the validation set does not.

During training, the CNN would tune the parameters of the convolutional filters so that they would extract the most pertinent features from the input images. These feature maps would then be used for classification / prediction. During validation, prediction performance on the validation set will be assessed, using the ground truth clinical data.

The specifications of the model training are as below:

- Framework: PyTorch
- CNNs: ResNet18 and DenseNet121
- Loss function: cross entropy loss
- Optimization algorithm: stochastic gradient descent (SGD)
    - Learning rate = 0.001 and Momentum = 0.9
- Number of epoch: 15

3. RESULTS

3.1 Validation Results For Each Condition

The validation results for different training conditions (CNN model, normalization method, slice selection method) are tabulated accordingly in Table 3.1.1 and Table 3.1.2 below. For each condition, the validation set has a total of 5 positives and 8 negatives in each fold.

Table 3.1.1: Validation Results Of ResNet CNN

Normalization	Slice	Fold	True Positives	True Negatives	False Positives	False Negatives
Regular	Mid	1	2	4	4	3
		2	2	7	1	3
		3	4	6	2	1
		4	3	6	2	2
Regular	Selected	1	4	3	5	1
		2	2	8	0	3
		3	3	5	3	2
		4	3	4	4	2
Custom	Mid	1	2	6	2	3
		2	2	5	3	3
		3	0	8	0	5
		4	3	8	0	2
Custom	Selected	1	2	7	1	3
		2	1	8	0	4
		3	2	7	1	3
		4	2	8	0	3

Table 3.1.2: Validation Results Of DenseNet CNN

Normalization	Slice	Fold	True Positives	True Negatives	False Positives	False Negatives
Regular	Mid	1	2	5	3	3
		2	2	8	0	3
		3	4	3	5	1
		4	2	5	3	3
Regular	Selected	1	2	5	3	3
		2	2	7	1	3
		3	4	2	6	1
		4	3	4	4	2
Custom	Mid	1	3	4	4	2
		2	3	8	0	2
		3	2	7	1	3
		4	4	4	4	1
Custom	Selected	1	2	7	1	3
		2	2	7	1	3
		3	2	7	1	3
		4	2	6	2	3

3.2 Sensitivity, Specificity And Accuracy Of Each Condition

The sensitivity, specificity and accuracy metric of each training condition are calculated, averaged over 4 folds and tabulated in the following Table 3.2.1.

Table 3.2.1: Calculated Sensitivity, Specificity And Accuracy Metrics

CNN	Normalization	Slice	Sensitivity	Specificity	Accuracy
ResNet	Regular	Mid	0.550	0.719	0.654
ResNet	Regular	Selected	0.600	0.625	0.615
ResNet	Custom	Mid	0.350	0.844	0.654
ResNet	Custom	Selected	0.350	0.938	0.712
DenseNet	Regular	Mid	0.500	0.656	0.596
DenseNet	Regular	Selected	0.550	0.563	0.558
DenseNet	Custom	Mid	0.600	0.719	0.673
DenseNet	Custom	Selected	0.400	0.844	0.673

This result is compared to a random chance classification. Given that both the training and validation set are not balanced with a positive to negative ratio of 5:8, a random chance model trained with this imbalance would have:

- expected sensitivity of 0.385
- expected specificity of 0.615
- expected accuracy of 0.527

3.3 Best Model Performance

The condition that yields the best performance is the one that uses ResNet CNN with custom normalization and manually selected slices. The average sensitivity, specificity and accuracy for this condition is plotted in Figure 3.3.1 below.

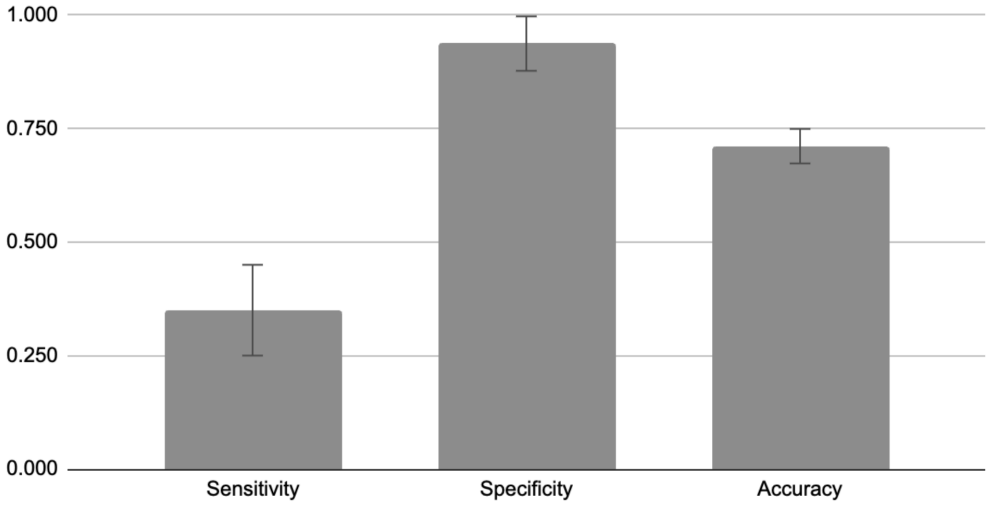

Figure 3.3.1: Best Model Performance Metrics

This best model performance has accuracy of 0.712 which is much better than chance of 0.527. This is confirmed using the Chi-square Test with the contingency table for the predicted and actual values as shown in Table 3.3.1 below.

Table 3.3.1: Contingency Table For Best Model Performance

	Predicted Positive	Predicted Negative
Actual Positive	7	13
Actual Negative	2	30

The two-tailed p-value equals 0.0017 (< 0.05), which implies that the observed frequencies are significantly different from what would be expected by chance. This confirms that the best model performance is indeed better than chance.

3.4 Comparison Between ResNet And DenseNet

The performance comparison between ResNet vs DenseNet in terms of sensitivity, specificity and accuracy is shown in Figure 3.4.1 below.

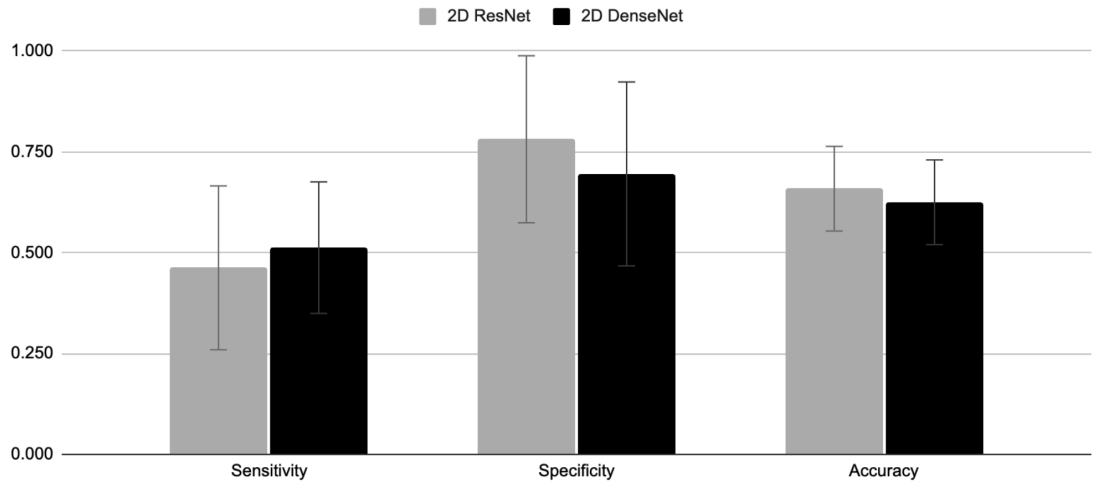

Figure 3.4.1: Performance Comparison Between 2D ResNet And 2D DenseNet

The contingency table showing the matching prediction results between the 2 CNNs is also shown in Table 3.4.1 below.

Table 3.4.1: Contingency Table Between 2D ResNet And 2D DenseNet

	DenseNet Correct	DenseNet Wrong
ResNet Correct	112	25
ResNet Wrong	18	53

The performance of 2 CNNs are comparable. Running McNemar's Test yields a two-tailed p-value of 0.36 (> 0.05), indicating that any difference in the performance between the 2 CNNs are not statistically significant.

3.5 Comparison Between Normalization Methods

The performance comparison between regular and custom normalization in terms of sensitivity, specificity and accuracy is shown in Figure 3.5.1 below.

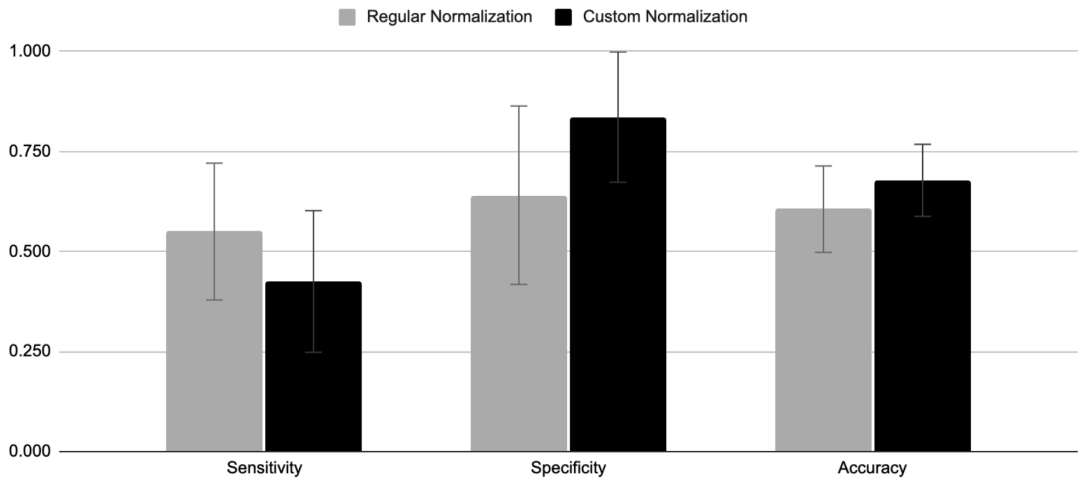

Figure 3.5.1: Performance Comparison Between Normalization Methods

The contingency table showing the matching prediction results between the normalization methods is also shown in Table 3.5.1 below.

Table 3.5.1: Contingency Table Between Normalization Methods

	Custom Correct	Custom Wrong
Regular Correct	110	16
Regular Wrong	31	51

The results show that the custom normalization performs better than the regular normalization. McNemar's Test yields a two-tailed p-value of 0.041 (< 0.05), indicating that this difference between the 2 normalization methods is statistically significant.

3.6 Comparison Between Slice Selection Methods

The performance comparison between mid slice and selected slice in terms of sensitivity, specificity and accuracy is shown in Figure 3.6.1 below.

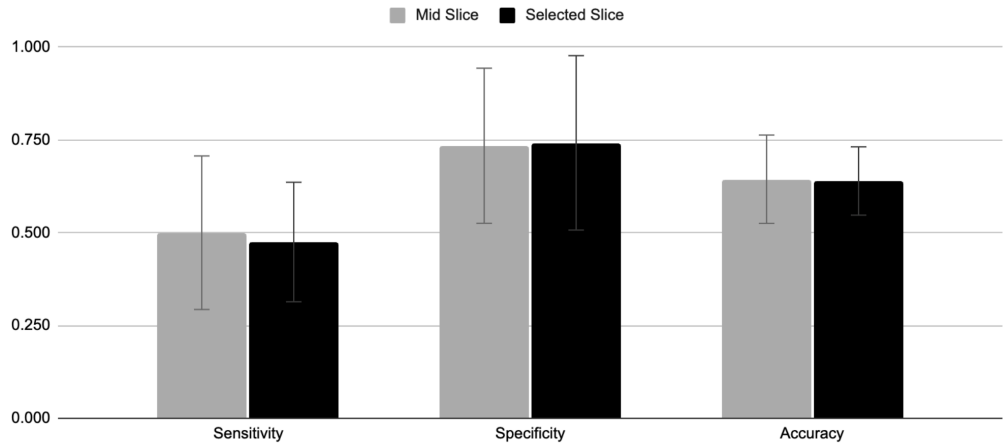

Figure 3.6.1: Performance Comparison Between Slice Selection Methods

The contingency table showing the matching prediction results between the slice selection methods is also shown in Table 3.6.1 below.

Table 3.6.1: Contingency Table Between Slice Selection Methods

	Custom Correct	Custom Wrong
Regular Correct	112	21
Regular Wrong	22	53

The performance of 2 slice selection methods are comparable. Running McNemar's Test yields a two-tailed p-value of 1.00 (> 0.05), indicating that any difference in the performance between the 2 slice selection methods are not statistically significant.

4. DISCUSSION

4.1 Best Model Performance

From Figure 3.3.1, the average sensitivity (0.350) indicates that the best model correctly identifies only about one third of the actual positive cases. This is just about the expected sensitivity of a random chance classifier, given the imbalance of our dataset (0.385), and may not be good enough for the use case of predicting positive cancer progression. We do not want to miss necessary therapies for the patients who have a high risk of progression. The average specificity (0.938) is notably much better than sensitivity and also better than the expected specificity of a random chance classifier (0.615). This shows that the model is more adept at correctly identifying negative cases and not recommending adjuvant therapies when the patients do not need them. The average accuracy (0.712) has been shown to be statistically better than the expected accuracy of a random chance classifier (0.527). Additionally, a naive model that always predicts the majority class would also only achieve an accuracy of approximately 0.615 (8/13), much lower than the best model performance. Coupled with the low standard deviation across the folds (0.038), the best model here has shown that its performance is consistent too.

Our results are in a similar range of overall accuracy as Lee et al., who also utilized CNNs to predict post-surgical progression and obtained overall accuracy of 0.77 in their larger cohort [8]. Park et al. predicted post-surgical progression using machine learning approaches (without CNNs), reaching an accuracy of 0.638, slightly lower than both our and Lee's group's accuracies [9]. Our results in the context of these others in the

literature confirm the utility of our CNN-based approach to predict biochemical recurrence post radical prostatectomy based upon the presurgical MRI data.

To further improve the model's performance in the future, a larger dataset is required to have more training data. This requirement is important for several reasons. Firstly, it would help improve generalization. A larger dataset can expose the model to a wider variety of examples, including edge cases and subtle variations. This helps the model learn the underlying nuanced patterns of the data and become better at generalizing to new, unseen data. Secondly, deep learning models with many parameters often require large amounts of data to prevent overfitting and to learn effectively. Hence, more training data enables the training of these complex architectures, potentially leading to improved performance.

Apart from that, the CNN architecture and hyperparameters can be further tuned to improve performance. We can experiment with different network depths, filter sizes, activation functions, optimizers and learning rates to yield better results.

4.2 Comparison Between ResNet And DenseNet

In Figure 3.4.1 and Table 3.4.1, ResNet and DenseNet are shown to have comparable performance and any difference between them is not statistically significant. Each network has their own strengths and limitations.

On the side of DenseNet, in terms of parameter efficiency, the dense connections allow the network to leverage features from earlier layers, promoting feature reuse and reducing the need for a large number of filters in each layer. This can be beneficial in the case where the dataset is limited in size. Furthermore, the direct connections between all layers in DenseNet facilitate better information and gradient flow throughout the network. This enables more efficient learning and can lead to better performance on complex classification tasks with subtle features in medical images.

On the other hand, ResNet has simpler architecture and is generally more straightforward. This simplicity can lead to reduced computational overhead and reduced feature redundancy. ResNet's additive skip connections create direct pathways for gradients to flow during backpropagation, preventing them from diminishing exponentially as they traverse deep layers. This helps address the vanishing gradient problem, enabling the training of much deeper networks with hundreds or even thousands of layers, unlocking greater representational power and performance gains.

Overall, the relative performance of ResNet and DenseNet depends heavily on the specific task at hand. More experiments with different cohorts and different ranges of cancer sizes are needed to determine the best network to use for this application.

4.3 Comparison Between Normalization Methods

From Figure 3.5.1 and Table 3.5.1, the custom normalization method performs statistically better than the regular min to max normalization method. This is indeed

expected. Due to the possible presence of artifacts, there may be many abnormally bright spots with maximum possible values in the image. If these large values are used as the scaling range max, most of the other values in the image would be much lower than this value, resulting in the abnormally dark image after normalization. In contrast, when we set the scaling range max to a lower value (the 80th percentile for T2W and a fixed upper value for ADC and DCE), this scaling operation would spread out all the pixel values more evenly across the range, resulting in an overall brighter image.

Figure 4.3.1 below shows an example where the custom normalization method results in a much brighter image than the regular normalization method. It enables finer details to be more easily seen and detected in the image with custom normalization.

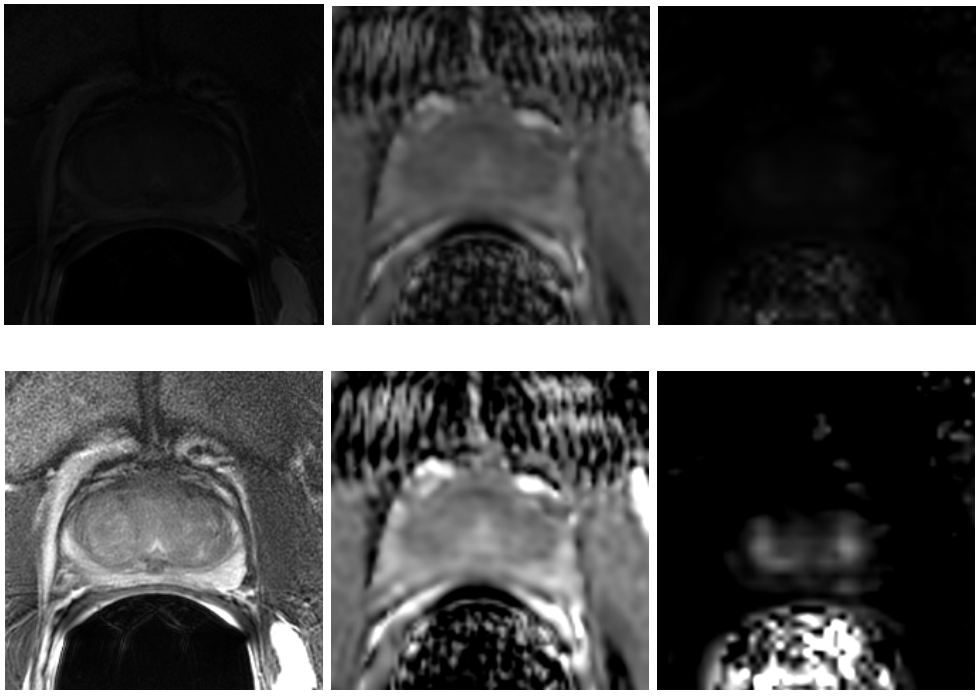

Figure 4.3.1: T2W, ADC, DCE After Regular (Top) And Custom (Bottom) Normalization

This custom normalization method has proven useful in the given dataset and it is expected to be important in other datasets too, especially those with more artifacts and less regular data. Moreover, there are other variations of custom normalization methods that can be explored and experimented with. For example, in the same fields of deep learning for prostate cancer research, instead of using the 0^{th} and 80^{th} percentile for min and max values, there was a proposal to use the 1^{st} and 99^{th} percentile values, which is expected to produce similar to our regular normalization method, and be inferior to custom normalization [10]. Another approach to normalization is that the values could be normalized to zero mean and unit standard deviation [11]. Future studies can experiment with these alternative approaches to find out the best one for the use case.

4.4 Comparison Between Slice Selection Methods

Figure 3.6.1 and Table 3.6.1 show quite a surprising result. Automatic mid slice selection and manually selected slice have comparable performance and any difference between them is not statistically significant.

The result is counterintuitive because selecting the slice with the most cancer area is expected to provide more information about the tumor for the CNN to learn. Sometimes, automatically selecting the mid slice may even miss the cancer area altogether.

Nevertheless, there may be benefits of selecting the mid slice too. The mid slice tends to be in the middle of the prostate and it may have the largest area of the prostate, capturing more information about both the healthy and the cancerous tissue.

Apart from the cancerous tissue, the healthy tissue may also contain subtle information about the cancer progression risk.

Furthermore, in our dataset, many times the mid slice is just the same slice as the manually selected one or just within a few adjacent slices. Especially when the cancer is large and spans multiple slides, it is possible that the mid slice can just show the cancer just as well as the selected slice. Figure 4.4.1 below illustrates this point with an example patient case. The mid-slice images show as much cancer information as the selected slice. The left peripheral zone is dark on the T2Ws and there is a small dark spot on both of the ADCs. The DCE slope is brighter in a larger area on the mid slice than the selected slice but they represent the cancer just the same.

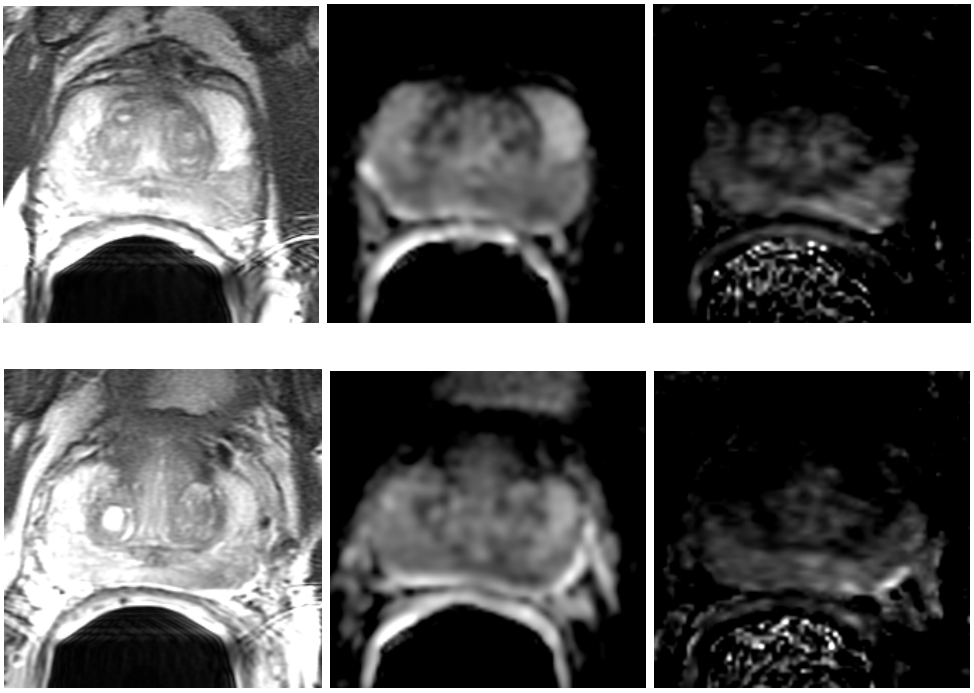

Figure 4.4.1: T2W, ADC, DCE From Mid Slice (Top) And Selected Slice (Bottom)

5. CONCLUSION

In this project, the application of convolutional neural networks (CNNs) in analyzing pre-surgical multiparametric MRI to predict prostate cancer progression has been explored. The results demonstrate the potential of CNNs as a valuable tool in this domain. The best-performing model, utilizing a ResNet architecture, custom normalization, and manually selected image slices, achieved an accuracy of 0.712, significantly surpassing random chance classification. The model's high specificity indicates its proficiency in accurately identifying negative cases, thereby avoiding unnecessary adjuvant therapies. However, the relatively lower sensitivity underscores the need for further refinement to enhance the model's ability to detect positive cases and ensure timely intervention for high-risk patients.

The comparative analysis between ResNet and DenseNet revealed comparable performance, suggesting that both architectures hold promise for this task. The choice between them may depend on specific dataset characteristics and computational constraints. The investigation into normalization methods highlighted the superior performance of the custom normalization technique, emphasizing the importance of addressing potential image artifacts for optimal model performance. The findings regarding slice selection methods, while somewhat counterintuitive, suggest that both automatic mid-slice selection and manual selection can yield valuable insights, potentially capturing different aspects of tumor characteristics and surrounding tissue.

The insights gained from this study pave the way for future research aimed at further enhancing the accuracy and clinical utility of CNN-based models for predicting prostate cancer progression. The acquisition of larger and more diverse datasets, coupled with improvements in model architecture and hyperparameter tuning, holds the potential to unlock even greater predictive power.

In addition, while 2D CNNs have shown promise, the inherent 3D nature of the prostate gland and its surrounding structures suggests that 3D CNNs could offer significant advantages and future studies should explore this further. 3D CNNs process the entire volumetric data, enabling them to capture spatial relationships and dependencies between adjacent slices that might be missed by 2D CNNs. This could lead to a more comprehensive understanding of tumor characteristics, such as shape, size, and location within the prostate, potentially improving the accuracy of cancer progression prediction.

All in all, the ongoing research and development of CNNs in the field of prostate cancer diagnosis and treatment holds immense potential to transform the way this disease is managed. By providing clinicians with more precise and efficient tools for image analysis, risk stratification, and treatment planning, CNNs can empower them to make more informed decisions that lead to improved patient outcomes. This technology has the potential to revolutionize prostate cancer management, ultimately leading to earlier detection, more personalized treatment plans, and increased survival rates.

REFERENCES

[1] Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74(1):12-49.

[2] Xia J, Trock BJ, Gulati R, et al. Overdetection of recurrence after radical prostatectomy: estimates based on patient and tumor characteristics. Clin Cancer Res. 2014;20(20):5302-10.

[3] Gleason DF. Histologic grading of prostate cancer: a perspective. Hum Pathol. 1992;23:273-9.

[4] Convolutional Neural Network. Wikipedia. Wikimedia Foundation. Updated May 24, 2022. Accessed August 28, 2024. https://en.wikipedia.org/wiki/Convolutional_neural_network.

[5] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:770-8.

[6] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017:4700-8.

[7] Gibbons M, Simko JP, Carroll PR, Noworolski SM. Prostate cancer lesion detection, volume quantification and high-grade cancer differentiation using cancer risk maps derived from multiparametric MRI with histopathology as the reference standard. Magn Reson Imaging. 2023;99:48-57.

[8] Lee HW, Kim E, Na I, et al. Novel multiparametric magnetic resonance imaging-based deep learning and clinical parameter integration for the prediction of

long-term biochemical recurrence-free survival in prostate cancer after radical prostatectomy. Cancers (Basel). 2023;15(13):3416. doi:10.3390/cancers15133416

[9] Park S, Byun J, Woo J. A machine learning approach to predict an early biochemical recurrence after a radical prostatectomy. Appl Sci. 2020;10(11):3854. doi:10.3390/app10113854

[10] Pellicer-Valero OJ, Marenco Jiménez JL, Gonzalez-Perez V, et al. Deep learning for fully automatic detection, segmentation, and Gleason grade estimation of prostate cancer in multiparametric magnetic resonance images. Sci Rep. 2022;12:2975. doi:10.1038/s41598-022-06730-6

[11] Saha A, Hosseinzadeh M, Huisman H. End-to-end prostate cancer detection in bpMRI via 3D CNNs: effects of attention mechanisms, clinical priori and decoupled false positive reduction. Med Image Anal. 2021;73:102161. doi:10.1016/j.media.2021.102161

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Duc Huy Doan
A6930708D69343A...
Author Signature

8/29/2024
Date