

ProstadeNet: A High Accuracy Fine-Tuned CNN Model for Prostate Cancer Grade Detection from Enhanced Histopathological Images

By
Sheikh Kaif
211-15-14682

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for
the **Degree of Bachelor of Science in Computer Science and
Engineering**

Supervised by

Ms. Nazmun Nessa Moon

Associate Professor

Department of Computer Science and
Engineering Daffodil International University

Co-Supervised by

Ms. Faiza Feroz

Lecturer

Department of Computer Science and
Engineering Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

May 14, 2025

APPROVAL

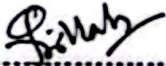
This Project titled "ProstadeNet: A High Accuracy Fine-Tuned CNN Model for Prostate Cancer Grade Detection from Enhanced histopathological images", submitted by Sheikh Kaif, ID: 211-15-14682 to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 14 May, 2025.

BOARD OF EXAMINERS



.....
Dr. Sheak Rashed Haider Noori
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Chairman



.....
Md Masum Billah
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



.....
Mr. Partha Dip Sarkar
Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



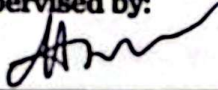
.....
Dr. Md. Zulfiker Mahmud
Professor
Department of Computer Science and Engineering
Jagannath University

External Examiner

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Ms. Nazmun Nessa Moon, Associate Professor**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:



Ms. Nazmun Nessa Moon

Associate Professor

Department of Computer Science and

Engineering Daffodil International University

Co-Supervised by:

Ms. Faiza Feroz

Lecturer

Department of Computer Science and

Engineering Daffodil International University

Submitted by:



Sheikh Kaif

Student ID: 211-15-14682

Department of Computer Science and

Engineering Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project(FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Ms. Nazmun Nessa Moon, Associate Professor**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Medical Image Processing** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Prostate cancer is the leading malignancy within the bodies of men globally and it is also the fifth most common cause of male cancer mortality. Prostate cancer needs to be diagnosed early in order to reduce mortality and give better treatment to concerned individuals. Although traditional methods of diagnosing prostate cancer are time-consuming, computer-aided diagnosis systems provide a quicker diagnosis without any loss of accuracy. There is a very alarming need for fast, accurate, and economical diagnostic processes of this epidemic. The main goal of this research is to recommend an effective framework for grading classification of prostate cancer from histopathological images with six classes with fewer errors of classification. Various image augmentation preprocessing techniques have been utilized for improving the quality of images, and four different augmentation techniques to enhance the dataset size and reduce model overfitting problems. Finally, a novel ProstadeNet architecture which is a tuned CNN tuned in structure and hyperparameters by ablation study will be presented. The ProstadeNet model achieved an accuracy of 99.83%. To estimate the capacity of ProstadeNet, relative experiments with five typical transfer learning models will be conducted to find out its relative performance. K-fold validation will be used in measuring results. Various performance measurements have been used to show the robustness of the models.

Keywords: prostate cancer; image preprocessing; CNN; transfer learning model, ablation study.

Table of Contents

Approval	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
List of Figures	vii
List of Tables	viii
1 Introduction	1
1.1 Introduction.....	1
1.2 Motivation.....	2
1.3 Objectives.....	3
1.4 Methodology.....	4
1.5 Project-Outcome.....	4
1.6 Organization of the Report.....	5
2 Background	6
2.1 Introduction.....	6
2.2 Literature Review.....	6
2.2.1 Related Research.....	11
2.3 Summary.....	12
3 Research Methodology	13
3.1 Methodology.....	13
3.1.1 Overview.....	13
3.1.2 Proposed Methodology.....	13

3.2	Detailed Methodology and Design.....	14
3.3	Task Allocation.....	19
4	Implementation and Results	20
4.1	Environmentt Setup.....	20
4.2	Testing and Evaluation.....	20
4.3	Results and Discussion.....	20
4.4	Summary.....	29
5	Engineering Standards and Design Challenges	30
5.1	Compliance with the Standards.....	30
5.1.1	Communication Standards.....	30
5.2	Impact on Society, Environment and Sustainability.....	30
5.2.1	Impact on Life.....	30
5.2.2	Impact on Society & Environment.....	31
5.2.3	Ethical Aspects.....	32
5.2.4	Sustainability Plan.....	33
5.3	Project-Management and Financial-Analysis.....	34
5.4	Complex Engineering Problem.....	35
5.4.1	Complex Problem Solving.....	35
5.4.1.1	Justification for EP Attributes Mapping.....	35
5.4.1.2	Justificatiomm for Knowledge Profile Mapping.....	36
5.4.2	Engineering Activities.....	37
5.4.2.1	Justification for Engineering Activities Mapping.....	37
6	Conclusion	38
6.1	Summary.....	38
6.2	Limitatiion.....	38
6.3	Future Work.....	38
	References	40

List of Figures

3.1 Overall strategy	13
3.2 The Dataset	14
3.3 Steps of Image pre-processing	16
3.4 Data Augmentation	17
3.5 The baseline model	17
3.6 Optimal CNN model based on Ablation study	18
4.1 The Model	20
4.2 The accuracy and loss curves	29
4.3 Result of 5-fold cross validation.	29

List of Tables

2.1 Literature Review	10
3.1 MSE, PSNR, SSIM, RMSE value for 5 images	16
3.2 Task Allocation	19
4.1 The accuracy and losses of all the transfer learning models	21
4.2 Case Study	22
4.3 Analysis of Best Model	27
5.1 Mapping with complex problem-solving	37
5.2 Mapping with knowledge Profile	37
5.3 Mapping with complex engineeringg activities	39

Chapter 1

Introduction

1.1 Introduction

Prostate cancer is the leading kind of cancer within the population of American males, prostate cancer as the third leading male cancer cause of death in 2017 with 161,360 new cases (19% of new male cancers) and 26,730 deaths (8% of deaths from cancer) [1]. Its occurrence is attributed to the overall slow growth of prostate cancer, with diagnosis early on accounting for survival [1]. Therefore, routine observation and diagnosis have to be conducted to augment survival of the patient. Clinical relevant diagnosis of PCa is currently based on a two-way interaction between PSA screening, DRE, TRUS, and MRI [2]. Moreover, PSA screening will lead to excessive rates of over-diagnosis and therefore redundant, expensive, and invasive needle biopsy and resultant over-treatment [2]. MRI with various parameters and, to a greater extent, DWI is now the imaging modality of choice, with radiologists able to report under ROC curves of 0.69-0.81 in PCa detection [3]. PI-RADS v2 system was also available for standardization of reporting [4], although inter-observer variability in its use is a concern [5].

Prostate cancer should be diagnosed in time so that it can be treated accordingly. The earlier it is diagnosed, the more and the less it will be in causing complications [6]. Prostate cancer detection and its analysis eats up a lot of time as well as being tiring because medical images have a lot of noises, artefacts and other complicated structures. Early diagnosis of prostate cancer is a challenge in some parts of the world because there is a lack of radiologists. There is a lack of radiologists by as much as 80% in some areas. These patients do not receive appropriate diagnosis and treatment for prostate cancer because of such unavailability of radiologists. Radiologists misdiagnose some patients of prostate cancer because they have no experience, no time, no workload, and too many patients.

In the clinical context, the CAD deep learning model is tasked with avoiding the effect of the radiologist shortage since it has the ability to diagnose the disease without any human intervention [7]. Prostate detection has been revolutionized by new dimensions in the form of medical imaging technology, deep learning, and computer vision. Artificial intelligence and machine learning techniques, i.e., crowds, received a humongous surge in the health sector, i.e., prostate [8] seeking. CNNs, as a deep learning technique, can be appropriate for detection of diseases many times after medical image processing and

histopathology image processing. Prostate cancer detection using CNNs can trigger a radiology face revolution and allow early detection of prostate cancer [9]. CNNs are superior to most systems in prostate cancer screening. Prostate cancer is detected better by radiologists, diagnosis is enhanced, and disease monitoring with PET/CT. As prostate cancer is a priority in global public health, additional research and activity should be performed to prevent, diagnose, and treat prostate cancer.

The objective of this exercise is to create a model that would classify images of prostate cancer histopathology into six classes in such a manner in which the rates of misclassification could be prevented.

There are a limited number of image enhancement methods employed for improving image quality and four types of data enhancement methods employed in creating the rest of the images. We have proposed ProstateNet with fine-tuned CNN and through augmentation with architectural and hyperparameter tuning after they are experimented with using ablation study. We compare ProstateNet with comparison of its output to five baseline transfer learning models. For the process of interpretation of results, use of a K-fold process will be done and several measures will be taken to make sure that the model is performing at its best level.

1.2 Motivation

The male population of United States have a higher possibility to develop prostate cancer. This cancer was responsible for approximately 8% of all cancer deaths and the third highest number of cancer deaths in men in the United States in 2017. Since prostate cancer is so common, being diagnosed early increases chances of survival, given its slow nature. Thus, early detection and prompt monitoring help patients live longer. Our present method of identifying significant prostate cancer (PCa) is through the use of PSA, DRE, TRUS and MRI [2]. Still, using PSA testing can occasionally result in picking up a problem that would not become serious, together with unnecessary, uncomfortable and expensive tests [2]. Nowadays, radiologists are choosing multiparametric MRI, with DWI at the forefront, to diagnose prostate cancer, as they can achieve an area under the ROC curve between 0.69 to 0.81 for finding PCa [3]. While PI-RADS v2 standardized image analysis, little consistency in its use is still seen [5]. The earlier prostate cancer is detected, the easier it is to manage. If the disease is found early, people have a greater chance to recover and are less likely to face problems [6]. Prostate cancer is not easy to detect and diagnose because many noises and other obstacles appear on medical images. In specific places around the world, it is hard to detect prostate cancer early since there are not enough specialized radiologists. In some locations, there could be a lack of radiologists by as much as 80%. A lack of radiologists in these regions may prevent many people from obtaining the appropriate care for prostate cancer. At times, mistakes in diagnosing prostate cancer are made by radiologists due to their lack of experience, rush, workload and overcrowded patients. The motivation to design this work comes from the need to detect and grade prostate

cancer earlier and more accurately, since it accounts for many men's cancer deaths. Diagnosis and accurate grading of cancer can still be delayed due to pathologist's need to review the low-quality histopathological images manually which can be slow, subjective and improper at times. If prostate cancer is found early, it affects the methods of treatment and influences how the disease develops in a patient. Yet, due to their complexity and the high number of data involved, it is not easy to accurately identify each case of prostate cancer. In this work, CNNs are used to automate the process of classifying prostate cancer into particular grades. With ProstateNet, the goal is to design a model that pathologists can use to find reliable results in less time and with greater ease. The motivation is also to decrease the number of deaths from prostate cancer by allowing doctors to catch it early and increase the accuracy of diagnosis with advanced processing and AI. In the end, this study intends to fill the space between the current methods for diagnosis and the use of CADs, as CADs support doctors and can help prevent death from cancer.

1.3 Objectives

The key primary objective of the study is to design a strong and effective deep learning-based framework for timely diagnosis of prostate cancer containing high classification accuracy based on the images of histopathology. Specifically, the study aims to:

Develop ProstateNet Model: Design and fine-tune a Convolutional Neural Network (CNN) model, ProstateNet, that perfectly segregates prostate cancer into six different grades from histopathological pictures.

Enhance Image Quality: Apply different histopathological image prepositions techniques like reduction of noise, contrast enhancement to enhance the quality of histopathological images for good model performance.

Evaluate Model Performance: Evaluate the performance of ProstateNet by comparison with five pre-trained transfer learning models on the basis of this model's capacity to outperform existing solutions with as concerns the accuracy and robustness.

Utilize K-Fold Cross-Validation: Exploit the K-fold cross-validation to make ProstateNet reliable and generalizable in order to avoid overfitting and improve the quality of the model for different datasets.

Measure Key Performance Metrics: Compare ProstateNet and other models based on various performance evaluation metrics such as- accuracy, precision, recall, and F1 score to find how effective both models and others are at detecting prostate cancer.

Establish a Cad System to Identify Early: Design and implement Computer-Aided Detection (CAD) system according to the ProstateNet model with a view of enhancing early diagnosis of prostate cancer hence lowering the rate of mortality and improving outcomes of patients.

This goal is the solution to a major void in the existing analysis of medical imaging and an attempt to develop the branch of cancer diagnostics.

1.4 Methodology

This research seeks to design ProstadeNet, a fine-tuned Convolutional Neural Network (CNN) model so as to achieve accurate prostate cancer grades classification from enhanced histopathological images. The research solves the problem of manual analysis of the images of prostate cancer, which is laborious and error-prone. Using a plethora of deep learning mechanisms cutting-edge image preprocessing approaches, the work focuses on increasing the diagnostic precision and assistance offered to pathologists in reaching more timely and reliable cancer grade denotations. The study will commence with the pre-processing of histopathological images using procedures such as noise reduction and contrast enhancement before the development and optimization of ProstadeNet using an ablation study. ProstadeNet will then be compared with 5 pre-trained transfer models in terms of relative performance. For the purpose of building robust model and making generalizations, K-fold cross-validation will be applied and various metrics including accuracy, precision and recall will be measured. The anticipated result will be the efficiency of ProstadeNet compared to other models in terms of prostate cancer grade classification with better accuracy and reliability. Overall, the study is aimed at creating a Computer-Aided Detection (CAD) system that may be implemented in clinical practice, since such usage would help identify prostate cancer early and lower mortality rates due to precise and timely diagnosis.

1.5 Project Outcome

The anticipated result of this study is that the ProstadeNet Convolutional Neural Network (CNN) model will successfully be developed and be able to classify prostate cancer grades using enhanced histopathological images. The task will allow us to prove that ProstadeNet is more accurate in diagnostics than the currently used pre-trained transfer learning models and deliver a worthwhile and effective solution for early diagnosis of prostate cancer. Key expected outcomes include:

Increased Accuracy in Gradation of Prostate Cancer: The ProstadeNet model is anticipated to have a high level of precision while putting prostate cancer into six levels with minor error considering its state-of-the-art and fine-tuned architecture in addition to its advanced prostate image pre-processing.

Enhanced Image Processing: The utilisation of diverse pre-processing techniques including noise reduction, enhancement of the images' contrast is expected to make the quality better of histopathological images, thus greatly contributing to the enhanced model performance.

Validation Through K-Fold Cross-Validation: The study is expected to prove that the use of K-fold cross-validation will guarantee the generalizability and robustness of the model, does not entail overfitting and validates the model performance with respect to various datasets.

Comparison with Pre-Trained Models: ProstadeNet is expected to perform better than five pre-trained transfer learning models when it comes to important performance parameters like accuracy, precision and recall making it superior for prostate cancer detection.

Potential for Clinical Integration: The research will lead to the development of a Computer-Aided Detection (CAD) system that can be incorporated into clinical workflows where the pathologists will be able to make faster and more accurate decisions that could potentially aid in the improvement of patients' outcomes and a decrease in mortality.

All in all, the scope of the study is to add a new and efficient solution to the sphere of medical image analysis, promoting timely detection of prostate cancer and improving the quality of the solution given to concerned individuals down the road.

1.6 Organization of the Report

This document is structured into several chapters, each addressing specific components of the research:

Chapter 1: Introduction presents a research overview, contextualizing the background, current conditions, problem formulation, aims, parameters, constraints, and a research summary.

Chapter 2: Literature Review thoroughly examines relevant scholarly works, conducts comparative analysis, highlights unresolved questions, and synthesizes key findings.

Chapter 3: Methodology/Requirement Analysis & Design Specification articulates the methodological approach, system architecture, hardware and software specifications, project administration techniques, and economic assessment.

Chapter 4: Implementation describes the development phases, encompassing model training procedures, prototype creation, system evaluation, and performance assessment.

Chapter 5: Result and Analysis exhibits experimental outcomes, simulation results, performance indicators, comparative evaluations, and their significance.

Chapter 6: Impact on Society, Environment, and Sustainability investigates the research's broader consequences for human existence, social structures, environmental considerations, ethical dimensions, and sustainability planning.

Chapter 7: Conclusion and Future Work offers research inferences, proposes subsequent research avenues, and addresses limitations and potential interest conflict

Chapter 2

Background

2.1 Introduction

In this section we are discussing in detail about the previous study, that was published based on prostate cancer grading detection. We selected the best paper that works with prostate cancer to detect cancer in the early stages.

2.2 Literature Review

Extensive work has been performed in determining or catching prostate cancer based on advanced artificial intelligence techniques, machine learning, and deep learning algorithms in the past two years [10,11,12].

Proposed a new model of prostate cancer diagnosis and lesion segmentation from MRI images with a variety of strategies including Support Vector Machines, Multilayer Perceptrons, and K-Nearest Neighbors with accuracy of 80.97% and 0.79 Dice coefficient.[13]. Created an algorithmic model using machine learning with some algorithms for prostate cancer identification such as Random Forest, Support Vector Machine, Deep Neural Network, Multilayer Perceptron, Logistic Regression, Linear Discriminant Analysis, and Linear Regression [14]. Among them, Multilayer Perceptron gave maximum accuracy of 97%, 100% recall, 0.958 AUC, 95% precision, and 97% F1-score. Compared fully connected Artificial Neural Networks, Random Forest, Support Vector Machine, and Logistic Regression models in predicting the prognosis of prostate cancer in 790 active surveillance patients, and Support Vector Machine classifier produced the highest F1-score of 0.586 [15]. Designed a deep learning approach to enhance the detection of prostate cancer based on various PSA molecular forms like free PSA, P2PSA, PSA density, and total PSA with a sensitivity of 86% and specificity of 89% in 437 patients [16]. Developed a novel deep learning technique for prostate cancer segmentation using batch normalization in 230 MRI images to improve feature extraction to capture multi-level features for prostate location and shape identification with 95.3% accuracy [17]. Used a deep learning model with pathology and radiology data embedded in it and matched 400 patients' MRI scan and biopsy with suspected prostate cancer for malignancy vs. benign labeling and risk level estimation with an AUC of 0.89 [18].

Used a Convolutional Neural Network in their deep learning model for the prognosis of lymph node metastasis of prostate cancer from tumor histology screening [19]. Their model achieved AUC 0.83 in 218 patients' stained tumor slides. Introduced a deep learning method for prostate cancer diagnosis and Gleason score grading of histopathological images for four datasets. Their post-processing improved CNN model achieved precision 0.8921 and F1-score 0.8460 [20]. Used deep learning techniques for prostate cancer detection from MRI images of 158 patients [21]. Performance with ultrasound-guided and MRI-guided biopsies at 82% accuracy, precision at 86%, and recall at 78% as well. StrengthsRegNet: deep learning registration

pipeline of prostate histopathology to MRI images. Their method, having trained on 654 image pairs of 152 prostate cancer patients, attained Dice coefficient scores between 0.96 and 0.98 [22]. Developed an independent tracking and detection system of implanted marker seeds in prostate cancer patients using deep learning. The markers facilitated the localization of the prostate volume and position during treatment. Their CNN detected 1,500 images at 98% accuracy [23].

Developed an auto-detection system of prostate cancer using multi-parametric MRI data for 780 patients. Their CNN achieved an extremely high AUC of 0.9684 [24]. Created a computerized diagnostic system for prostate cancer based on deep learning. Compared 10,616 histology slides across various CNN architecture and feature extraction techniques and attained the highest accuracy rate of 92.77%. Created a model to predict prostate cancer from 330 mpMRI images of the Prostate X challenge dataset [25]. Their DenseNet and MobileNet-based comparison of prostate screening tactics attained AUC of 0.931 and F1-score of 0.93 [26]. Employed a deep learning network to identify prostate cancer from MRI images with a Dice score of 90.89% in 85 patients [27]. They also performed detection of prostate cancer on the biopsy slides, and it consisted of 5,759 prostate biopsies on 1,243 slides. Their CNN had an outstanding AUC of 0.992 [28].

Author (s)	Year	Title	Methodology	Key Findings
Si Nae Oh	2024	DL-CAD + ENS for Pneumothorax	Retrospective cohort study using difference-in-differences analysis.	CAD+ENS significantly reduced time to oxygen therapy; no significant effect on other treatments.
Tauhidul Islam	2024	Review of DL Data Augmentation in Medical Imaging	Systematic review of recent literature, datasets, algorithms, and evaluation metrics.	DL-based augmentation of data increases medical model quality; highlights trends, challenges, and future directions.
Noémie Rabilloud	2023	Review of DL in PCa Pathology	Systematic review of 77 studies using PubMed and Embase; bias assessed with adapted QUADAS-2 tool.	DL shows strong performance in PCa diagnosis (AUC up to 0.99); promising but limited predictive use due to lack of validation.
Mingze He	2023	DL in MRI for PCa Diagnosis	Comprehensive literature review examining deep learning implementations in MRI-based prostate cancer diagnosis and therapeutic approaches.	DL shows promise in tumor detection, lesion segmentation, and prognosis prediction; accuracy can improve with more data and model optimization..

Madhur Nayan	2022	ML for PCa Progression Prediction	Retrospective cohort study; evaluated ML models (SVM, random forest, ANN, ML-LR) and T-LR for predicting grade-progression.	ML models significantly outperformed traditional logistic regression in predicting PCa progression on active surveillance.
Mehmet Emin Salman	2022	YOLO-based PCa Detection System	YOLO object detection model fine-tuned on annotated prostate biopsy images; tested on two sets.	Achieved 97% accuracy on similar test set, 89% on unseen set; effective at detecting and grading prostate cancer.
Keya Amarsee	2021	DL for Tracking Fiducial Markers in PCa	Convolutional neural network YOLO v2 employed for fiducial marker seed detection in kilovoltage X-ray imaging.	Deep learning detected fiducial marker seeds in 98% of images with ≤ 1 mm variation from ground truth.
Albert Comelli	2021	DL for Prostate MRI Segmentation	Applied UNet, ENet, and ERFNet for prostate gland segmentation on T2-weighted MRI with k-fold validation and Tversky loss.	ENet achieved 90.89% DSC and faster performance (6s per image) compared to UNet, showing potential for real-time prostate delineation.
Hans Pinckaers	2021	End-to-End WSI Prostate Cancer Detection	Used a streaming CNN implementation (ResNet-34) to train on whole-slide biopsy images (268MP) with only image-level labels, avoiding manual pixel-wise annotations.	Achieved AUC of 0.992, comparable to MIL methods, while using full-slide context; streaming is slower but allows training on large datasets without detailed labels.
Keya Amarsee	2021	Deep learning algorithm application for automatic identification and monitoring of marker seeds	Trained YOLO v2 on corrected kV images for fiducial detection.	98% detection rate with ≤ 1 mm error; enables accurate real-time tracking.

		implanted in prostate cancer patients.		
Francesco Gentile	2021	Enhanced identification methodology for high-grade prostate cancer utilizing various PSA molecular forms combined with PSA density	Neural net trained on PSA metrics and age to predict Gleason scores.	86% sensitivity, 89% specificity for high-grade cancer; fewer unnecessary biopsies.
Li Zhang	2021	Novel MRI methodology for prostate cancer analysing.	Segmented MRI using GrowCut, extracted features, and classified with ensemble models.	~20% accuracy gain in lesion detection over prior methods.
Frederik Wessels	2021	Histology-based deep learning framework for direct prediction of lymph node metastasis from primary tumours in prostate cancer.	CNN trained on H&E slides to predict lymph node metastasis.	AUROC 0.76–0.83; better than existing clinical models.
Octavian Sabin Tătaru	2021	Contemporary trends and future directions in artificial intelligence and machine learning applications for prostate cancer patient management.	Reviewed AI across pathology, imaging, genomics, and therapy.	AI improves detection, personalization, and precision surgery; validation needed.
Ali Hasan Md. Linkon	2021	Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: An extensive study	Surveyed DL-based CAD for grading via whole-slide images.	CNNs outperform traditional tools but face reproducibility and data hurdles.
Ebru Erdem	2021	Comparative analysis of diverse supervised machine learning techniques for prostate cancer prediction.	Compared 9 ML models using patient data for cancer prediction.	MLP achieved 97% accuracy and top AUC/F1 scores.

Abhishek Patel	2021	Deep learning framework implementation for prostate cancer detection.	Tested ML (SVM/CNN) with SIFT, EFD, and mp-MRI features.	Combined models reached AUC ~0.90; beat PI-RADS.
Wei Shao	2021	ProsRegNet: Deep learning architecture designed for registration between MRI and histopathology prostate images.	Built ProsRegNet to align histology with pre-op MRI.	20× faster, more accurate cancer label transfer for training.
Pegah Khosravi	2021	Pathology-radiology fusion approach utilizing deep learning for diagnostic classification of prostate cancer.	AI-biopsy DNN trained on MRI and pathology scores.	Outperforms PI-RADS for risk detection without manual input.

Table 2.1: Literature Review.

2.2.1 Related Research

The most common cancer among American men is prostate cancer. It was the third highest cause of cancer-related death in men in 2017 with 161,360 new cases (19% of new cancer) and 26,730 deaths (8% of all cancer death) [1]. Since the prostate cancer does grow at a relatively slow rate, time does allow survival rates to skyrocket if diagnosed in its early years [1]. Active diagnosis and detection are therefore mainly important to maximize patient survival. PSA screening, DRE, TRUS, and MRI are currently the gold standard for clinically significant prostate cancer (PCa) diagnosis [2]. Yet, over-diagnosis is quite frequent with PSA screening and results in needless, expensive, invasive needle biopsies and perhaps over-treatment [2]. Multiparametric MRI, and more specifically DWI, are becoming ever more utilized as the imaging study of choice, and radiologists are reporting an area under the ROC curve of 0.69 to 0.81 for the detection of PCa [3]. PI-RADS v2 system was created with the intent of standardizing image interpretation [4], but observer variability remains an issue on application [5].

Early detection is the gateway to treatment and recovery of prostate cancer. The sooner the detection, the higher the percentage of recovery and the lower the chance of having complications [6]. Prostate cancer detection and diagnosis is a time-consuming process because there are a large number of noises, artefacts and other complicated structures in medical images. Early detection of prostate cancer still remains challenging in some parts of the world since the world is running extremely short of radiologists. Shortage of radiologists in some parts of the world is up to 80%. Due to such shortage of radiologists, patients in such parts of the world may not get their rightful diagnosis and treatment of prostate cancer. Radio-diagnostic mistakes in the diagnosis of prostate cancer are made because of the lack of skills, time constraint, workload enhancement and patient overload in the radiology departments.

In medical imaging of healthcare, deep learning-based CAD system takes the responsibility of minimizing the effect of a lack of radiologists by automatic reporting of the existence of the disease [7]. Recent progresses in deep learning, computer vision, and medical imaging technology over the last few years have improved opportunities for prostate identification. Use of machine-learning and artificial intelligence (AI) techniques, particularly use of CNN has also been miraculous in medicine and beyond this in prostate detection [8]. CNNs are one of the deep learning techniques in which medical images such as histopathological image data can be used for disease identification effectively. Use of CNN for prostate cancer detection can revolutionize the discipline of radiology and take the detection of prostate cancer grade at an early stage to a new height [9]. The CNNs identified the prostate cancer more precisely compared to the conventional image analysis techniques. It would help the radiologists in the diagnosis of prostate cancer and help them in diagnosing and treating the disease more precisely and accurately. Prostate cancer, being the most severe public health issue worldwide, needs more research and therapy, diagnosis, and prevention. The task is to develop an efficient prostate cancer grading system based on histopathology images divided into six classes with fewer errors.

Certain Image preprocessing techniques are to be done for quality improvement of the image. We will present our new model, ProstateNet, a design of a fine-tuned CNN based model, and architectural and hyperparameter modifications on the model are guided by the result of ablation study. To compare the performance of ProstateNet and five transfer learning pre-trained models in comparative effectiveness estimation, it is compared. It is compared separately for each outcome with K-fold validation, and different performance measures are compared in comparing the robustness of the model.

2.3 Summary

In conclusion, although advancement has been achieved towards harnessing the usage of AI and deep learning in prostate-cancer grading detection, there is need to continue with research and innovations to overcome the current challenges.. The implementation of cutting-edge technologies may contribute to improving the accurateness of diagnosing, decreasing the level of subjectiveness, and providing early detection in a more affordable and convenient manner. Such efforts are crucial in enhancing patient outcomes and improving the global burden related to prostate cancer, especially amongst resource-limited areas.

Chapter 3

Research Methodology

3.1 Methodology

3.1.1 Overview

In this section, we review the proposed methodology and system design, where we explain the corresponding hardware and software requirements that are necessary in implementing the system. The area of project management defines the planning, performance, and tracking of the activities of the project for it to be timely and successfully implemented. In addition to this, we offer a financial analysis to describe budget, cost and possible economic benefits of the project. In the end, we summarize the main points, insisting on the value of the suggested approach to improving the prostate cancer diagnosis process and the ability to apply high-tech approaches in clinical practices.

3.1.2 Proposed Methodology

In this work, we propose an innovative approach aimed at solving the problems of prostate cancer detection specifically. The approach undertaken in this study incorporates image preprocessing, CNN model, transfer learning model and an ablation study for improved accuracy and efficiency for detection of prostate-cancer in the various classes. Figure 3.1 shows the overall process here.

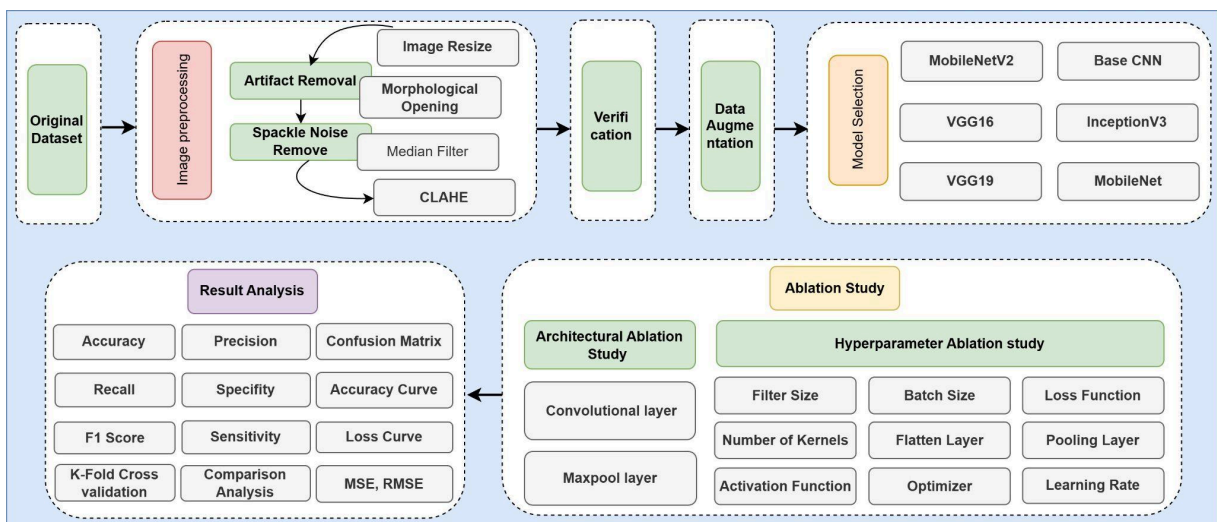


Figure 3.1 Methodology

3.2 Detailed Methodology and Design

Dataset Description

This study used an prostate cancer dataset from online including 10,616 prostate images, each labeled with a severity grade ranging from zero to five, indicating whether the cancer is significant or non-significant [29].

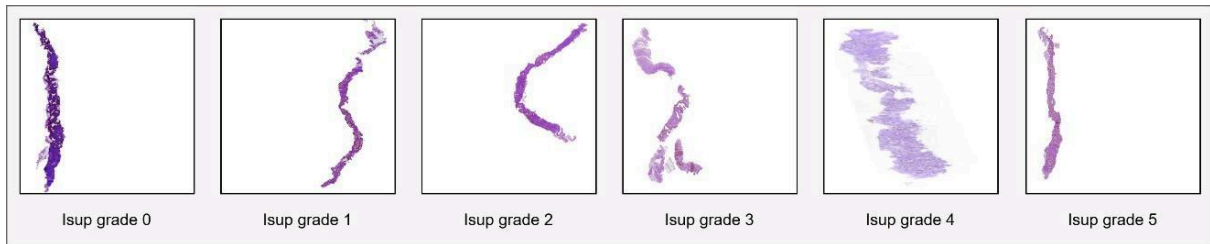


Figure 3.2 The Dataset

Image pre-processing

The prostate cancer grading dataset includes a lot of noise and artefacts. In order to improve the accuracy of the model, we employ various image preprocessing techniques to reduce the amount of noise and artefacts in the dataset. Image processing is the initial stage for feeding the DL-model since images are often inclusive of noise and artefacts, which must be removed before the model can be trained. Before Spackle noise removal we resized the images into 224 x 224. A morphological opening was performed to clean up this image and remove the artefacts. Before that, a median filter was used to get rid of the noise. CLAHE was employed for image enhancement to better feature understanding of the images.

Image Resize

Resizing adjusts an image's dimensions, usually to fit a desired aspect ratio or specific dimensions, while maintaining or changing its visual quality [30]. The images are resized into 224 x 224 from 512 x 512.

Speckle Noise Removal

The prostate cancer grading dataset has a significant amount of speckle noise. The use of median filters allows for the elimination of speckling noise. Due to this, the median filter is first applied in dataset to produce images free of speckle noise [31]. The expression of the speckle noise can be expressed like equation 1:

$$S(A, B) = M(A, B) N(A, B) \quad (1)$$

Where $S(A, B)$ is the images with spackle noise, $M(A, B)$ is the original images and $N(A, B)$ is the speckle noise

Median Filter

Histopathological images are usually characterized by noise of different forms including that of spackle and as such requires noise reduction to maintain diagnostic authenticity. A median filter can be used in the processing of histopathological images to efficiently suppress this noise, without serious damages of details in the images [32] making the images clearer and fidelity. It is quite helpful in cases of cancer diagnosis applications where precise image analysis is of importance.

Morphological Opening

Images tend to present many small objects, edges, or blobs following noise reduction, and such a problem has to be addressed to achieve maximum accuracy in deep learning models. In histopathology pictures, the morphological opening operations were employed to highlight the feature contours and successfully eliminate small entities [33]. It performs operations of dilation and erosion that may improve the contrast, and get rid of the noise while the image is analyzed; thus, increasing the precision of the analysis. Equation 1 gives the expression of the morphological opening [34]:

$$A \circ C = (A \ominus C) \oplus A \quad (1)$$

Where the input image is A , the structuring element is S , \ominus and \oplus represents the erosion and dilation respectively.

CLAHE

Contrast Limited Adaptive Histogram Equalization (CLAHE) is a superior technique applied to enhance local contrast in an image which is mainly effective in complex structures in medical imaging. Unlike the conventional adaptive histogram equalization, CLAHE uses contrast limiting to avoid the amplification of the noise at the expense of finer details. This approach improves local contrast in a selective way and considerably increases the readability of medical images, which makes it possible to see intricate anatomical structures more clearly [36].

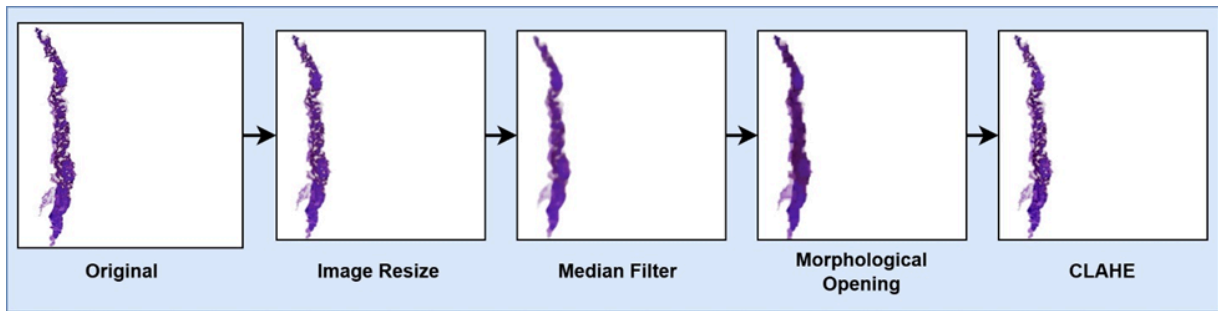


Fig 3.3 Steps of Image pre-processing

Verification

Quality of processed image can be determined by different algorithms. The following work measures quality and the application of a set of established verification metrics, which are the MSE [27], RMSE [28], PSNR [29], and SSIM [30]. These metrics are commonly used in the assessment of preprocessing techniques on image integrity preservation. The below table illustrates the MSE, RMSE, PSNR, and SSIM between the five captured images, and is a depiction of the comparison of the images' quality.

Image	MSE	PSNR	SSIM	RMSE
Image_1	12.88	38.35	0.95	0.12
Image_2	14.23	36.79	0.95	0.13
Image_3	15.49	44.36	0.96	0.11
Image_4	14.38	36.36	0.94	0.12
Image_5	14.35	39.28	0.96	0.11

Table 3.1 MSE, PSNR, SSIM, RMSE value for 5 images

Table 2 presents the evaluation of image quality using verification metrics such as MSE, PSNR, SSIM, and RMSE for five processed images. The results confirm that the applied image processing techniques did not degrade image quality, ensuring the preservation of essential visual information.

Data Augmentation

After the image processing, a number of photometric augmentations were made, such as: changing the tone, brightness, contrast, and focus. These modifications help create wide range of visual variations, thus rendering the dataset more robust. By simulating different lighting and model generalization, these augmentations enhance the imaging condition [36]. This ensures improved performance in everyday situations due to reduced sensitivity for certain visual features. Once the augmentation are applied, the amount of images on the dataset increased to 53,080.

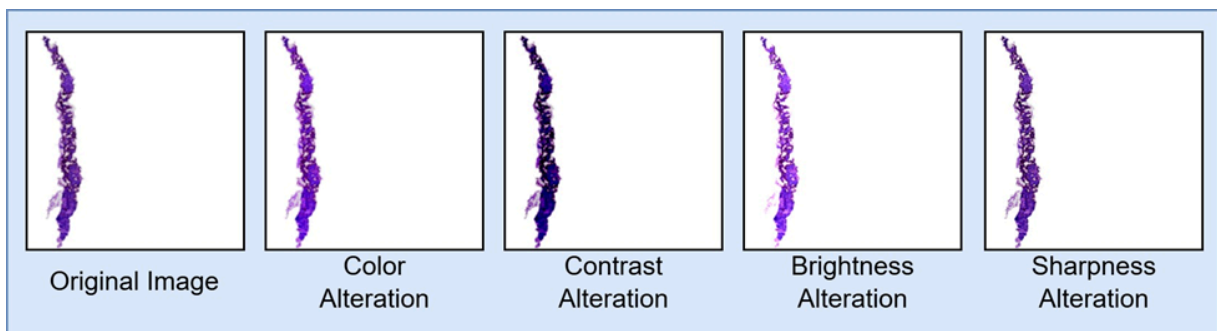


Fig 3.4 Data Augmentation

Base CNN Model

This experiment start, by implementing a baseline CNN model architecture developed from the scratch for purposes of grading prostate cancer. The model has five convolutional. three-layered arrangement and five max-pooling with form of three blocks. The first block contains one the second block consists of two which consists of two convolutional layers with the subsequent max-pooling layer, and the third block consists of two followed by a single max-pooling-layer, the convolutional-layers. After these layers, the architecture has a flattening-layer, a fully connected layer, and an output layer, a SoftMax activation function. The highest layer uses PReLU function as the activation function, followed by SoftMax. The loss function is the categorical cross-entropy and the Nadam is being used optimizer is applied with learning rate of 0.0001. At the beginning, the training process was performed with a batch. size 32, that was later upgraded to 64 for better operation. Figure 3.5 illustrates the complete architecture of the baseline model.

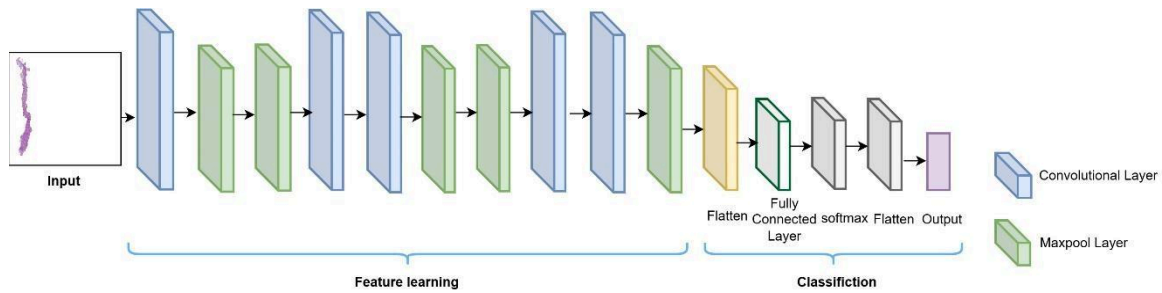


Figure 3.5 The baseline model

The CNN model discussed above is based on two integral elements: feature learning and classification. The convolutional-max pooling layers perform in the feature learning phase in order to extract meaningful patterns out of the input data. The classification phase then passes these features that were obtained through extraction through a chain of layers such as a flattening layer to transform the data, a dense layer for the subsequent computation by shoulders of the neural, and a dropout-layer with a 0.5. dropout rate to prevent overfitting. This is followed by a SoftMax layer, SoftMax layer is used to generate probability distributions over classes, and a last dense dense output layer for classification. This architecture is set to process and classify data efficiently, hence making CNNs highly suitable for image recognition and a host of similar tasks. In doing this, the model integrates these components puts in place strong feature extraction and prediction capabilities.

ProstadeNet

Based on the ablation studies, this study suggests a shallow, robust, and resilient model architecture, LungNet-13. LungNet-13 has three convolutional layers, and a maxpool layer follows each convolutional layer. In all the convolutional layers, 3 x 3 convolutional kernels were utilized to the model efficacy. The ReLU activation function enhances the model performance with a stride size of 1 x 1. Softmax was used as an activation function with a 0.001 learning rate to further finetuned the model performance. A dropout rate of 0.5 was also used to prevent the mmodel architecture's overfitting.

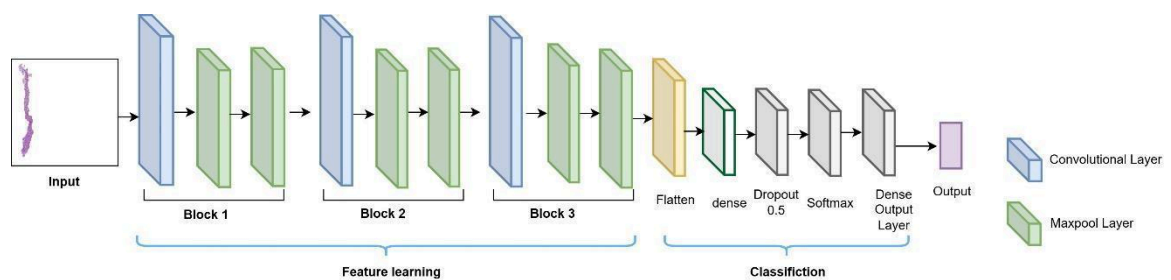


Figure 3.6 Optimal CNN model based on Ablation study

Figure 3.6 illustrates the optimized and robust CNN model, refined through an architectural study. The proposed ProstadeNet consists of two primary components: a feature learning module, responsible for extracting meaningful patterns from input data, and a classification module, which processes these features to generate accurate predictions.

3.3 Task Allocation

Task	Timeline	Allocated to
Data collection	Week 1-4	Me
Data Preprocessing	Week 5-6	Me
Model Development	Week 7-9	Me
Evaluation	Week 10-13	Me
Documentation	Week 13-14	Me
Deployment Planning	Week 15	Me

Table 3.2 Task Allocation

Chapter 4

Implementation and Results

4.1 Environment Setup

Setup involves using Python 3.7+ with essential libraries like TensorFlow, Keras, NumPy, and Matplotlib, preferably on Google Colab with GPU support enabled for faster training. Required packages can be installed via pip, and datasets should be organized into structured folders for training and validation. Google Drive integration is recommended for saving models and logs using ModelCheckpoint and CSVLogger. This setup makes sure smooth model development, training, and evaluation. Python 3.6.6 is used for Jupyter Notebook.

4.2 Testing and Evaluation

At this stage, experts use unseen data to measure the model's abilities after it has been trained. Success is measured by accuracy, precision, recall or F1-score, depending on what the task is. Existing data is tested with a distinct test set and the model's answers are compared to actual truths to compute these metrics. The goal of this step is to make sure that the trained model is able to handle real-life situations. You will find the model diagram in section 4.1.

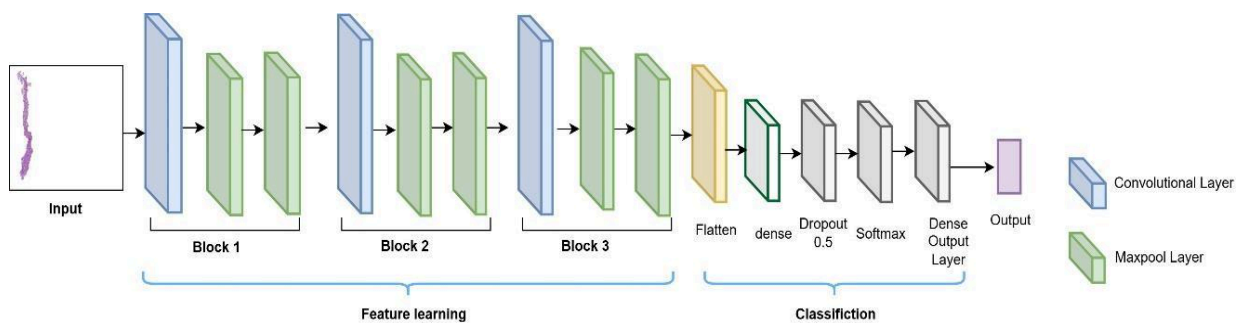


Figure 4.1: The Model

4.3 Results and Discussion

This investigation utilized various metrics to evaluate transfer learning models, comprising precision, recall, F1-score, ACC, sensitivity, and specificity. Confusion matrices were generated for each model, from which TP, TN, FP, and FN instances were derived. Area under curve (AUC) values were calculated in certain cases. Additionally, the statistical analysis of models included computation of FPR, false FNR, FDR, MAE, RMSE.

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$ACC = 2 \frac{precision*recall}{precision+recall} \quad (5)$$

$$FPR = \frac{FP}{FP+TN} \quad (6)$$

$$FNR = \frac{FN}{FN+TP} \quad (7)$$

$$FDR = \frac{FP}{TP+FP} \quad (8)$$

$$FPR = 1 - Specificity \quad (9)$$

$$MAE = \frac{1}{n} \sum_{j=1}^n (y_j - y_j^p)^2 \quad (10)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - y_j^p)^2} \quad (11)$$

Table 3 presents an overview of the accuracy metrics for training, test, and validation phases, alongside corresponding loss values for the models and the baseline CNN model. The data indicates that the baseline CNN model exhibited enhanced performance relative to the other evaluated models.

Table 4.1 The accuracy and losses of all the transfer learning models

Model	Train_Accuracy	Train_Loss	Val_Accuracy	Val_loss	Test_Accuracy	Test_loss
VGG-19	95.56	0.24	95.24	0.23	95.24	0.24
VGG-16	96.46	0.19	95.85	0.13	95.85	0.19
Mobile-Net	95.43	0.18	94.23	0.28	94.24	0.33
Mobile-Net V2	95.76	0.25	94.63	0.32	94.59	0.39
Inception-V3	90.86	0.421	90.21	0.392	90.21	0.384
Base CNN	96.98	0.21	96.07	0.22	96.07	0.23

The table below captures the five transfer learning models' performance and that of the baseline CNN-model. The VGG19 model, in particular, recorded accuracy of 95.56% and had both validation and test accuracies of 95.24%. It performed better than VGG16, which recorded a training of 96.46% and validation accuracy of 95.85%. MobileNet model obtained training accuracy of 95.43% and validation accuracies of 94.23% and 94.24%, respectively. MobileNetV2, on the other hand, obtained small improvements, with training accuracy of 95.76%, validation accuracy of 94.63%, and test accuracy of 94.59%. Its accuracy was lowest at training, validation and testing levels, where it obtained 90.86%, 90.21% and 90.21% accuracy respectively. On all metrics, the baseline CNN performed best with the highest training accuracy of 96.98%, while also having high validation and test accuracies of 96.07%. Why the base CNN performed best was the fact that it had the optimal results. The base CNN was thus used for future research. The network was then further improved by performing an ablation study

Ablation Study:

Here, we present the results of ten case studies in the ablation study. First, the professionals performed architectural ablation to pick the right model structure and then they adjusted the hyperparameters to boost the model's performance.

Table 4.2 Case Study

Architectural Ablation Study					
Case_Study-1: Changing Convolutional and Maxpool Layer					
No	Convolution layer	Pooling layer	Training time and epochs	Test accuracy	Findings
1	2	2	10s x 97	90.23%	Accuracy dropped
2	5	5	10s x 60	96.07%	Previous accuracy
3	3	6	9s x 60	96.83%	Highest accuracy
4	7	3	10s x 98	95.98%	Accuracy dropped
Hyperparameter Ablation Study					

Case_Study-2: Changing the kernel size				
No	Kernel size	Training time and epochs	Test accuracy	Findings
1	2 x 2	11s x 70	92.71%	Accuracy dropped
2	3 x 3	9s x 57	96.91%	Highest accuracy
3	4 x 4	11s x 71	89.99%	Accuracy dropped
4	5 x 5	9s x 60	96.83%	Previous accuracy
Case_Study-3: Changing the number of kernels				
No	Number of kernels	Training time and epochs	Test accuracy	Findings
1	16 x 32 x 64 x 128	9s x 69	94.98%	Accuracy dropped
2	64 x 64 x 64 x 64	9s x 60	96.91%	Previous accuracy
3	32 x 64 x 128 x 256	9s x 60	96.98%	Highest accuracy
4	32 x 64 x 64 x 128	10s x 71	94.99%	Accuracy dropped
1	16 x 32 x 64 x 128	10s x 299	90.96%	Accuracy dropped
Case_Study-4: Changing the pooling layers				
No	Pooling layer	Training time and epochs	Test accuracy	Findings
1	Max	9s x 60	96.98%	Highest accuracy
2	Average	10s x 65	96.82%	Accuracy dropped

Case_Stud-5: Changing the activation function				
No	Activation	Training time and epochs	Test accuracy	Findings
1	PReLU	9s x 60	96.98%	Previous accuracy
2	Leaky ReLU	10s x 60	93.99%	Accuracy Increase
3	ReLU	8s x 57	97.51%	Highest accuracy
4	ELU	11s x 85	91.23%	Accuracy dropped
Case_Study-6: Changing the flattened layer				
Configuration No.	Flatten layer types	Training time x epochs	Test accuracy	Findings
1	Flatten	8s x 57	97.51%	Highest accuracy
2	Global Max pooling	9s x 60	95.24%	Accuracy dropped
3	Global Average pooling	9s x 54	96.03%	Accuracy dropped
Case_Study-7: Changing the batch size				
Configuration No.	Batch size	Training time x epochs	Test accuracy	Finding
1	32	7s x 56	96.83%	Modest accuracy
2	16	6s x 50	97.93%	Highest accuracy
3	64	8s x 57	93.65%	Previous accuracy
4	128	9s x 60	92.06%	Modest accuracy

Case_Study-8: Changing the loss function				
Configuration No.	Loss Functions	Training time x epochs	Test accuracy	Findings
1	Binary Crossentropys	Error	Error	Error
2	Categorical Crossentropys	6s x 50	97.93%	Highest accuracy
3	Mean Squared Errors	8s x 62	96.82%	Accuracy dropped
4	Mean absolute errors	8s x 61	68.25%	Accuracy dropped
5	Mean squared logarithmic error	8s x 67	96.83%	Accuracy dropped
Case_Study-9: Changing optimizers				
Configuration No.	Optimizers	Training time x epochs	Test accuracy	Findings
1	Adam	5s x 45	98.41%	Highest accuracy
2	Nadam	6s x 50	97.93%	Previous accuracy
3	SGD	7s x 62	84.13%	Accuracy dropped
4	Adamax	7s x 61	90.48%	Accuracy dropped
Case_Study-10: Changing learning rate				
Configuration No.	Learning rates	Training time x epochs	Test accuracy	Findings
1	0.0001	5s x 52	98.41%	Previous accuracy

2	0.001	5s x 45	99.83%	Highest accuracy
3	0.01	5s x 50	98.28%	Accuracy dropped

The process is ablation displayed in table above. By altering the number of convolutional and max-pooling layers, it became simpler to increase performance and sturdiness. By using 3 convolutional layers and 6 max-pooling layers, we managed to get 96.83%. Unfortunately, the smallest computed accuracy of 90.23% was found when the model was configured with 2 convolutional and 2 max-pooling layers. It was found that a kernel size of 3x3 offers the highest accuracy rate. The highest accuracy of 96.98% was achieved when using max-pooling over average pooling. The ReLU activation function produced the best results, achieving 97.5% accuracy in just 8 seconds. Additionally, the flattened layer configuration led to the highest accuracy, while the optimal batch size was 16, resulting in a 97.93% accuracy when using categorical cross-entropy. When comparing optimizers, Adam outperformed others, such as Nadam, SGD, and Adamax, with a test accuracy of 98.41%. Finally, the optimal learning rate of 0.001 was found to produce the highest accuracy of 99.83%.

Performance Analysis of Best Model:

The precision of the classification is significantly improved after employing best architecture as well as hyperparameters for the introduced ProstadeNet model. Table 4.3 lists a detailed overview of the best model configuration, identifying all the significant parameters and settings for the introduced structure. The structured presentation makes the model simpler to understand and reproduce, ensuring clarity and transparency in the model design.

Configuration	Value
Image sizes	224 x 224

Pooling layer	maxpool
Kernel size	3 x 3
Number of Kernel	32 x 64 x 128 x 256
Optimization	Adam
Activation	ReLU
Learning-rate	0.001
Batch size	16
Stride size	1 x 1
Flattened layer	Flatten
Activation-functions	Softmax
Dropout	0.5
Momentum	0.9
Accuracy	99.83%

Table 4.3 Analysis of Best Model

As indicated in the above table, this research utilized 224 x 224-pixel images as input to the model. A layer has had a 3 x 3 kernel size, and ReLU was the activation function for training model. Adam-optimizer (0.001) is used. The batch-size was 16, and Softmax-activation was used in the output layer. To prevent overfitting, dropout of 0.5 was utilized, and a momentum of 0.9 was utilized during training. The model achieved a final accuracy of 99.83%.

Statistical Analysis of the Proposed Model:

Following performance evaluation of model in question, statistical testing is conducted to ensure its strength and reliability. FPR quantifies the proportion of times negative samples are incorrectly classified, while FNR approximates the proportion of positive samples undetected.

FDR calculates the proportion of false positives over optimistic predictions. KC is a measure of evaluation used to compute the concordance between observed and predicted values. Meanwhile, MCC assists in measuring the correlation of real and predicted classification values. MAE and RMSE were also implemented.

Accuracy	FPR (%)	FNR (%)	FDR (%)	KC (%)	MCC (%)	MAE	RMSE	Precession	Recall	Specificity	F1 Score
99.83	0.001	0.003	0.003	1.00	1.00	0.00	0.00	99.69	99.69	99.93	99.69

Accuracy was superb with 99.83%. While the false negative rate was only 0.001% and the false discovery rate was only 0.003%, the false positive rate was maintained at a very minimum of 0.003%. The Kappa coefficient of the model was 1.00%, and the Matthews correlation value was 1.00%, indicating that the predicted and observed class labels were in great agreement. The RMSE and MAE were both 0, indicating low inaccuracy in the model's predictions. The model achieved a very high specificity of 99.93%, a very high recall of 99.69%, and a very high precision of 99.69%. The F1 score of 99.69% also indicated outstanding performance. Accuracy and loss plots are visual model performance evaluation utilities. The accuracy plot tracks classification performance over time and epochs, and the loss plot shows the model's loss function statistics. Together, they are utilized to assess model performance and look for potential overfitting.

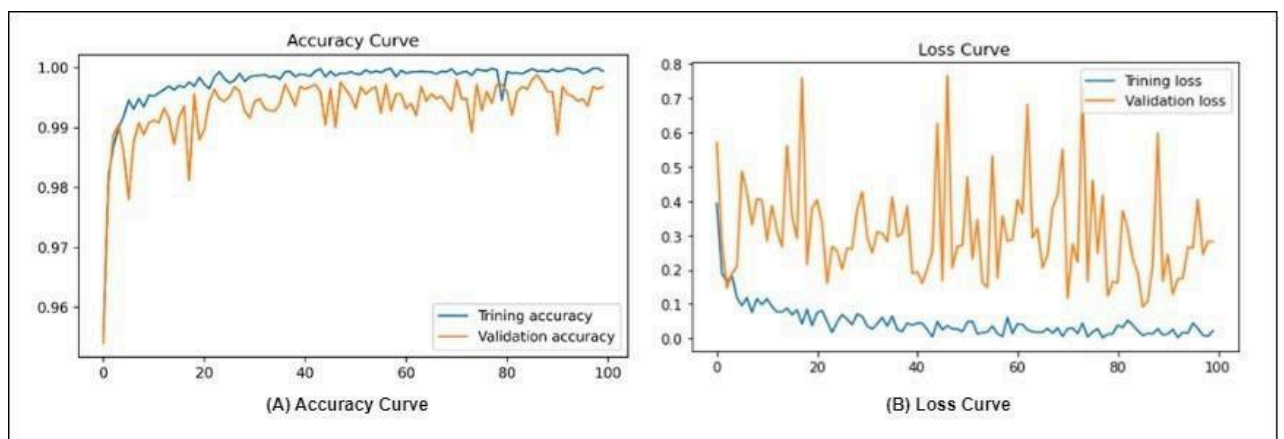


Figure 4.2 The accuracy and loss curves.

K-fold Cross Validation:

K-fold-cross-validation allows a model's effectiveness to be evaluated in machine learning [43]. Using this method, you disperse data into K-parts, train on K-1, test it on the last one and repeat the process K times [44]. For this part, the study applies 5-fold-cross-validation, where the dataset is separated into five data subsets, each used for a test run and the rest used to train the model. Figure 4.3 exhibits the five-fold-cross-validation results to assess if the proposed model is overfitting.

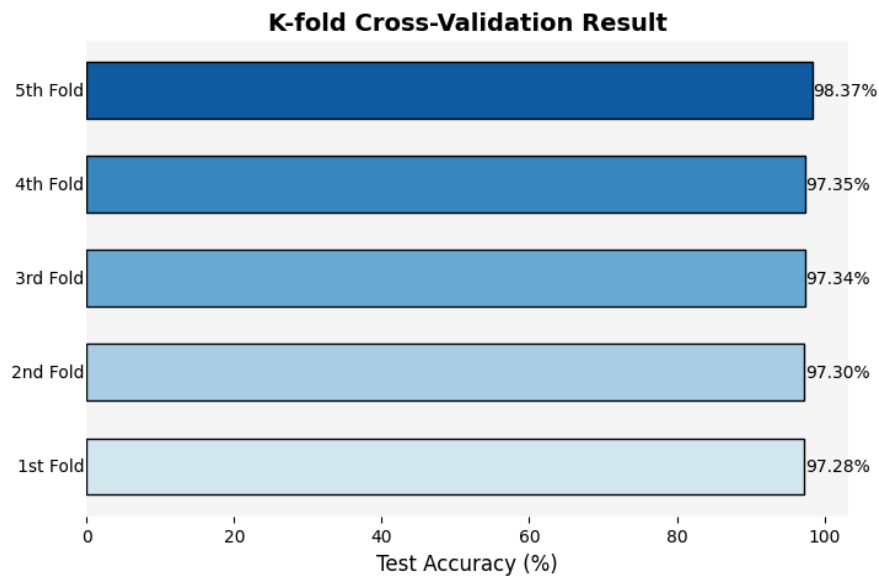


Figure 4.3 Result of 5-fold cross validation.

4.4 Summary

This study involves an in-depth exploration of trainable parameters by iteratively adjusting the model architecture and fine-tuning parameters using training data. The goal is to optimize the neural network's configuration, validate its effectiveness. Through systematic experimentation, we aim to enhance the model's performance, achieving better accuracy and generalization in tasks such as prostate cancer detection.

Chapter 5

Engineering Standards and Design Challenges

5.1 Compliance with the Standards

5.1.1 Communication Standards

During the initial phase of this study, efforts were made to collaborate with medical professionals for collecting real-world prostate cancer histopathological data. Communication was established with specialists, including those affiliated with Dhaka Medical College. However, due to administrative restrictions, ethical approval complexities, and limited accessibility to medical records, acquiring original clinical data proved to be extremely challenging. As a result, publicly available and well-curated datasets were utilized for model development and validation. While real clinical data would have further enriched the study, the selected dataset still allowed for a reliable and reproducible evaluation of methodology.

5.2 Impact on Society, Environment and Sustainability

In this section we discussed impact on society, impact on environment, limitations and future-work.

5.2.1 Impact on Life

The scope of application of this study to human existence is extensive, particularly with regard to the early detection of prostate cancer and improved patient care. Prostate-cancer is one of the most common cancers in men, and classifying it correctly helps determine the optimal treatment. Traditional diagnostic methods, i.e., clinical manual histopathologic diagnosis, are time-consuming, subjective, and susceptible to observer variation. With ProstateNet employing a highly optimized CNN structure, we can now automatically differentiate prostate cancer cases more accurately. With this innovation, medical misdiagnosis of cancers, delayed treatment, and waiting time to appropriate care may be minimized. If the deep learning model is applied to a CAD system, then to an extent, it can alleviate pathologists' workload because they can perform test hard cases as initial evaluation can be done with the help of AI software. Moreover, through the application of

image-preprocessing methods. The authors are capable of presenting histopathological images that have been improved to a greater extent for proper analysis. With this technology at hand, medical experts will be positioned to diagnose disease at an early stage, thus making the treatment more effective. In turn, it will help evade death from prostate cancer alongside quality care. The use of automated devices can also be beneficial to hospitals where there are no, or poor, pathologists. Through this research, a new means should hopefully be made available for the application of AI and to show diagnostic procedures especially in cancer diagnosis.

5.2.2 Impact on Society & Environment

The implications of this study in the society is immense in that it helps in the medical advancement technology, facilitating better healthcare accessibility, and finally leading to a lower burden on prostate-cancer. A prostate cancer is one of the major public health issues, and the diagnostic accuracy of it is crucial has direct implications on treatment successes and survival rates. By introducing ProstateNet, a deep prostate cancer grade learning-based Computer-Aided Detection (CAD) system is a garbling classification, this research agrees with the incorporation of artificial intelligence in the modern world healthcare making more diagnostic tests efficient and reliable. One of the major societal benefits lies in the chances of eliminating healthcare disparities, especially among low-resource units. neighborhoods with poor access to the experienced pathologists and superior diagnosis facilities. An AI-based diagnostic system can normalize the grading of cancer, cutting inter-observer variation. variability and making sure that all patients are diagnosed similarly and correctly. in terms of geographical location or institutional expertise. Furthermore, an aspect which ensures that there is a minimal need for redundant biopsies and minimising the overdiagnosis resulting from implementation of the PSA based screening, this study may contribute to reducing the medical expenditure and physical and mental suffering. with invasive procedures. The incorporation of AI in healthcare also promotes interdisciplinary. coordination of medical experts and data scientists, to foster innovation of improving overall medical research. Besides, when AI-powered tools become more commonly accepted, they can change the perception of people about the role the technology can play in medicine. growing trust towards automated healthcare solutions and the further development AI-driven diagnostics. In the long term, this study adds up to the future in which early and precise diagnosis of such conditions as prostate cancer becomes easier, affordable and indeed, useful, resulting in improved patients' outcomes, fewer deaths, and a healthier society overall. The influence of

this study on the environment is primarily indirect but still rather high as it is essential in developing sustainable and efficient practices in healthcare. Traditional physical resources, such as the use of the biopsy are heavily relied upon in order to diagnose the prostate cancer. procedures, chemical staining agents, and laboratories testing, which all produce os are biomedical waste and utilize a lot of energy and material. By developing ProstateNet; — an AI-based CAD for prostate-cancer grade classification, this study will be promoting an efficient and less resource consuming diagnostic process. Medical diagnostics will benefit from the integration of deep learning models, as this will curtail the process dependence on the physical samples by enabling the digital histopathological images to be analyzed to a high level of accuracy reducing the need for further biopsy and lab tests. This decrease in invasive measures results in the decrease in the amount of medical waste generated, including biological dangerous materials, chemicals, and one time only medical supplies. Furthermore, by through automating diagnostic process, AI-driven systems reduce the dependence on physical. infrastructure like big scale laboratories and manually operated diagnostic equipments which typically have high energy consumption. Cloud-based AI models can also optimize resources as well. application through processing medical data efficiently without demanding one to have a higher computational power power, most preferably optimized for the sustainable machine learning practices. Additionally, as The amount of the patient's travel to the special medical centers is decreased thanks to AI-based diagnostics, there is a possible decrease of carbon emissions that could be related to the transportation, contributing to a greener healthcare system. This study aims at achieving efficiency, accuracy, and sustainability coincides with the global trend to the ecologically responsible advances in medicine proving that technological developments in healthcare can also show good results implications for environmental conservation.

5.2.3 Ethical Aspects

This study addresses several ethical considerations related to the development and implementation of an AI-driven CAD for prostate-cancer grade classification.

Confidentiality and Security of Patient Data: Since the images used are from histopathology, confidentiality is immensely needed. In order to avoid private medical information reaching unauthorized parties, data anonymization and compliance with laws such as HIPAA and GDPR are necessary. Abnormal distribution of data in AI models can cause AI to name some diseases incorrectly for people who belong to such categories. The data of healthcare must provide diversity in patients so that the model does not discriminate against patient characteristics. Reports given by AI are precise,

but physicians must always keep their context in mind while treating patients. If prostate cancer is incorrectly classified, the doctor treating the patient may choose a less suitable treatment, and the patient may develop health problems. To achieve good and reliable results, algorithms must be thoroughly vetted, continuously monitored and expert views must be incorporated. All stakeholders involved, patients and healthcare providers, must be aware of how AI operates in the provision of medical care, what it can and cannot do and where it draws the line. Translucency and explainability of models make everyone trust and accept the results generated by AI.

Responsibility and Accountability: When AI-assisted treatment mistakes happen, it is important to define who is responsible. It is necessary to confirm who ought to take responsibility for keeping the system dependable and for addressing any negative consequences.

An Effect on Doctors and Nurses: While AI may speed up diagnosis, it raises ethical concerns for those working in the medical field. AI tools ought to be used in such a way that they help pathologists, without interfering with the important human skills and knowledge needed in medical diagnosis.

By addressing these ethical aspects, developing AI-based diagnostic system that is not only technologically advanced but also responsible, fair, and aligned with the highest standards of medical ethics.

5.2.4 Sustainability Plan

Long-term impact and viability of ProstateNet, the sustainability plan focuses on key aspects such as technological scalability, financial feasibility, environmental responsibility, and continuous improvement.

Efficiency and Range: The model will ensure that it can be used on many devices, including computers, cloud, remote online systems and in-hospital gadgets. To keep the model reliable, regularly update it, train it using fresh data and improve it with explainable AI methods.

Cost-Effectiveness: The goal is to design the system so that healthcare institutions, having limited resources, can easily adopt it. To ensure its steady growth, the organization can rely on open-source collaborations, partnerships with healthcare organizations and money from research grants or government help. Offering AI as a cloud service on subscription could make it easier for hospitals to afford the technology.

Environmental Protection: Keeping the environmental impact of diagnosing prostate cancer low by avoiding accidents resulting in more biopsy sessions and reducing the use of unnecessary laboratory materials and computational power. By prioritizing energy-efficient technology and cloud computing, less energy will be consumed. Furthermore, thanks to digital analysis, there is a reduced need for physical samples which helps decrease medical wastes. To be useful in practical situations, ProstateNet ought to be included smoothly in workflows already used by medical staff. It should be

adjustable to use with software for running hospitals, managing telemedicine and devices that work with AI. For easy adoption and effective usability, it will be important to partner with healthcare institutions. For AI in diagnostics to last, it needs people who have received training, can interpret the outcomes, fix problems and make necessary improvements to the model. If all medical practitioners, pathologists and radiologists regularly train, they will be able to perform their jobs effectively. New educational tools and programs on AI will be offered to help everyone better understand and trust AI-related medical diagnoses. Regular research will be done to maintain ProstateNet's performance, identify ways it is biased and ensure it remains clear and easy to understand. Prostate cancer diagnosis can be improved further by working closely with AI developers, doctors and public health professionals.

By implementing this sustainability plan, ProstateNet can achieve long-term success as a reliable, cost-effective, and environmentally responsible solution for prostate cancer detection, ensuring it remains beneficial for patients, healthcare providers, and society as a whole.

5.3 Project Management and Financial Analysis

As far as project management goes, my supervision directed and monitored the development process, making sure that deadlines and project objectives were met. In order to accomplish the intended results, I had to buy a GPU that costs 50,000 BDT for running the model efficiently, consuming very less time. I also needed to buy a Printer for printing my report which costs around 7000 BDT. And 1500 BDT was spent as fare while consulting doctors regarding my dataset.

5.4 Complex Engineering Problem

In this section, we discussed complex engineering problem about the research project “ProstadeNet: A High Accuracy Fine-Tuned CNN Model for Prostate Cancer Grade Detection from Enhanced histopathological images”. It explores the boundaries of medical imaging and machine learning, necessitating the creation of innovative computational techniques, the management of extremely complicated, frequent, and advanced analytical abilities.

5.4.1 Complex Problem Solving

Table_5.1: Mapping-with-complex-problem-solving.

EP-1 Dept of Knowledge	EP-2 Range Of Conflicting Requirements	EP-3 Depth of Analysis	EP-4 Familiarity of Issues	EP-5 Extent of Applicable Codes	EP-6 Extent Of Stake- holder Involvement	EP-7 Interdepend ence
✓		✓	✓		✓	

Mapping with Knowledge Profile for EP1

This table 5.2 is designed to map the EP1 to the Knowledge Profile.

Table_5.2: Mapping-with-knowledge-Profile.

K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K8 Research Literature
✓	✓		✓	✓

5.4.1.1 Justification for EP Attributes Mapping

EP1: Depth of knowledge

Strong knowledge of deep learning, medical image processing, and model optimization through techniques like ablation study and architecture tuning.

EP3: Depth of Analysis

Extensive analysis includes preprocessing, augmentation, cross-validation, performance comparison with five models, and ablation studies for model improvement.

EP4: Familiarity of Issues

The study addresses a well-known clinical issues like prostate-cancer detection using familiar deep learning techniques but introduces a novel model for improved performance.

EP6: Extent of Stake- holder Involvement

Though academic, the system has high clinical relevance and can benefit doctors and patients in real-world prostate cancer diagnosis.

5.4.1.2 Justification for Knowledge Profile Mapping (linked to EP1):

K3: Engineering Fundamentals

Applies to core engineering principles, such as image processing algorithms, convolution operations, and classification techniques essential in computer vision tasks.

K4: Specialist Knowledge

It demonstrates advanced expertise in deep learning, histopathological image analysis, CNN architecture design, and medical image classification.

K6: Engineering Practice

Practical skills were applied in model implementation using modern tools (TensorFlow, OpenCV), along with techniques like augmentation and validation to ensure robustness.

K8: Research Literature

The study is grounded in existing research on prostate cancer diagnosis and deep learning, using comparative evaluation with transfer learning models from the literature.

5.4.2 Engineering Activities

Table_5.3: Mapping-with-complex-engineering-activities.

E-A1 Range of re- sources	EA-2 Level of Interaction	EA-3 Innovation	EA-4 Consequences for society and environment	EA-5 Familiarity
	✓	✓	✓	✓

5.4.2.1 Justification for Engineering Activities Mapping:

EA2: Level of Interaction

It involves interaction between multiple domains like engineering, healthcare, and computer science for ensuring that technical methods align with clinical needs.

EA3: Innovation

The proposed ProstateNet model introduces a novel CNN design with fine-tuned hyperparameters and improved accuracy over existing transfer learning models.

EA4: Consequences for society and environment

Positive societal impact by potentially enabling early and accurate prostate cancer detection, which can improve survival rates and reduce healthcare burdens.

EA5: Familiarity

The work is based on well-established practices in image classification and medical AI, while extending them through original model design and advanced validation techniques.

Chapter 6

Conclusion

6.1 Summary

Early-detection and proper diagnosis are both essential for successful treatment and prevention of prostate cancer. In spite of the paucity of radiologists, computer vision and deep learning technologies can potentially greatly improve prostate cancer grading and alleviate the global burden of the disease. This paper presents a CNN-based model that predicts six specific grades from ISUP Grade 0 through 5 with a remarkable accuracy of 99.83%. Several image-processing-techniques are used for pre-processing the histopathological images, and several photometric augmentations are used to increase the instances. The model classifies data initially, and then the performance of the model is assessed by an ablation and it compares the models of transfer learning and the proposed model. Overfitting is avoided using K-fold cross-validation. This method brings in a new opportunity for early-detection and staging of prostate-cancer, guiding clinicians in their decision-making.

6.2 Limitation

How well ProstateNet does depends on the quality, range and generality of the images in the histopathological image data set. Low diversity of the data set may cause the model to do poorly on other patient groups. It is also worth mentioning that deep learning models need disproportionate computational power, and it becomes difficult to implement them to limited clinical setups. As those models are hard to interpret, some medical doctors may be hesitant to trust their output. In addition, because ProstateNet is not fully validated and standardized in everyday practice yet, large-scale clinical use is not possible. Ethical and regulatory issues, along with possible overdependence on AI, call for cautious and responsible integration into clinical routines.

6.3 Future Work

In further work, the dataset will be enlarged and made more diverse improving the model's robustness and fairness. In terms of clinical trust, enhancing interpretability with the help of explainable AI techniques (e.g., attention mechanisms, feature attribution techniques) will be essential. Experimentation will also target in making the model computationally efficient, to

allow real-time analysis in resource-constrained conditions by means of light architectures or edge AI implementation. Future clinical validation in several institutions will allow for the adaptation to the different imaging protocols. Moreover, inclusion of all evaluation metrics – for instance, confusion matrices, AUC (area under the curve), and ROC (receiver operating curve) curves– will provide performance insights. Long-term objectives would involve alignment to regulatory standards, integration with multi-modal diagnostic applications, and finalization of a CAD system that would provide pathologists with the help in the diagnostic pipeline.

References

- [1] Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2017. *CA: a cancer journal for clinicians* 67, 7–30 (2017).
- [2] Sandhu, G. S. & Andriole, G. L. Overdiagnosis of prostate cancer. *Journal of the National Cancer Institute Monographs* 2012, 146–151 (2012).
- [3] Sonn, G.A.et al. Prostate magnetic resonance imaging interpretation varies substantially across radiologists. *Eur. urology focus* (2017).
- [4] Hassanzadeh, E.et al. Prostate Imaging Reporting and Data System Version 2 (PIRADS v2): A pictorial review. *Abdom Radiol* 42, 278–289 <https://doi.org/10.1016/j.trsl.2014.08.005>. (2017).
- [5] Rosenkrantz, A. B. et al. Interobserver reproducibility of the pi-rads version 2 lexicon: a multicenter study of six experienced prostate radiologists. *Radiology* 280, 793–804 (2016).
- [6] Scardino, P.T., Weaver, R. and M'Liss, A.H., 1992. Early detection of prostate cancer. *Human pathology*, 23(3), pp.211-222.
- [7] Oh, S.N., Yang, H., Lee, C.K., Park, S.H., Han, C.H., Heo, H. and Kim, Y.S., 2024. Impact of Deep Learning-Based Computer-Aided Detection and Electronic Notification System for Pneumothorax on Time to Treatment: Clinical Implementation. *Journal of the American College of Radiology*.
- [8] Tătaru, O.S., Vartolomei, M.D., Rassweiler, J.J., Virgil, O., Lucarelli, G., Porpiglia, F., Amparore, D., Manfredi, M., Carrieri, G., Falagario, U. and Terracciano, D., 2021. Artificial intelligence and machine learning in prostate cancer patient management—current trends and future perspectives. *Diagnostics*, 11(2), p.354.
- [9] Yoo, S., Gujrathi, I., Haider, M.A. and Khalvati, F., 2019. Prostate cancer detection using deep convolutional neural networks. *Scientific reports*, 9(1), p.19518.
- [10] Salman ME, Çakar GC, Azimjonov J, Kösem M, Cedimoğlu IH, (2022) Automated prostate cancer grading and diagnosis system using deep learning-based Yolo object detection algorithm. *Expert Syst Appl* 201:117148
- [11] Rabilloud N et al (2023) Deep learning methodologies applied to digital pathology in prostate cancer: a systematic review. *Diagnostics* 13(16):2676
- [12] He M et al (2023) Research progress on deep learning in magnetic resonance imaging-based diagnosis and treatment of prostate cancer: a review on the current status and perspectives. *Front Oncol* 13:1189370
- [13] Zhang L, Li L, Tang M, Huan Y, Zhang X, Zhe X (2021) A new approach to diagnosing prostate cancer through magnetic resonance imaging. *Alex Eng J* 60(1):897–904
- [14] Erdem E, Bozkurt F (2021) A comparison of various supervised machine learning techniques for prostate cancer prediction. *Avrupa Bilim ve Teknoloji Dergisi* 21:610–620
- [15] Nayan N et al. (2022) "A machine learning approach to predict progression on active surveillance for prostate cancer." vol. 40: Elsevier, 4 ed., pp. 161–e1
- [16] Gentile F et al (2021) Optimized identification of high-grade prostate cancer by combining different PSA molecular forms and PSA density in a deep learning model. *Diagnostics* 11(2):335
- [17] Shrestha S, Alsadoon A, Prasad PWC, Seher I, Alsadoon OH (2021) A novel solution of using deep learning for prostate cancer segmentation: enhanced batch normalization. *Multimed Tools Appl* 80(14):21293–21313

- [18] Khosravi P et al (2021) A deep learning approach to diagnostic classification of prostate cancer using pathology–radiology fusion. *J Magn Reson Imaging* 54(2):462–471
- [19] Wessels F et al (2021) Deep learning approach to predict lymph node metastasis directly from primary tumour histology in prostate cancer. *BJU Int* 128(3):352–360
- [20] Linkon AHM, Labib MM, Hasan T, Hossain M (2021) Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: an extensive study. *Inform Med Unlocked* 24:100582
- [21] Patel A, Singh SK, Khamparia A (2021) Detection of prostate cancer using deep learning framework. In: *InIOP Conference Series: Materials Science and Engineering* (Vol. 1022(1), p. 012073). IOP Publishing
- [22] Shao W et al (2021) ProsRegNet: a deep learning framework for registration of MRI and histopathology images of the prostate. *Med Image Anal* 68:101919
- [23] Amarsee K et al (2021) Automatic detection and tracking of marker seeds implanted in prostate cancer patients using a deep learning algorithm. *J Med Phys* 46(2):80
- [24] Yang H, Wu G, Shen D, and Liao S (2021) "Automatic prostate cancer detection on multi-parametric mri with hierarchical weakly supervised learning". *IEEE* pp. 316–319
- [25] Kovalev VA, Voynov DM, Malyshau VD, Lapo ED (2020) Computerized diagnosis of prostate cancer based on whole slide histology images and deep learning methods. *InInformatics* 17(4):48–60
- [26] John J, Ravikumar A, Abraham B (2021) Prostate cancer prediction from multiple pretrained computer vision model. *Heal Technol* 11(5):1003–1011
- [27] Comelli A et al (2021) Deep learning-based methods for prostate segmentation in magnetic resonance imaging. *Appl Sci* 11(2):782
- [28] Pinckaers H, Bulten W, van der Laak J, Litjens G (2021) Detection of prostate cancer in whole-slide images through end-to-end training with image-level labels. *IEEE Trans Med Imaging* 40(7):1817–1826
- [29] Balaha, H.M., Shaban, A.O., El-Gendy, E.M. et al. Prostate cancer grading framework based on deep transfer learning and Aquila optimizer. *Neural Comput & Applic* 36, 7877–7902 (2024). <https://doi.org/10.1007/s00521-024-09499-z>
- [30] Talebi, H. and Milanfar, P., 2021. Learning to resize images for computer vision tasks. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 497-506).
- [31] Mohan, E., Rajesh, A., Sunitha, G., Konduru, R.M., Avanija, J. and Ganesh Babu, L., 2021. A deep neural network learning-based speckle noise removal technique for enhancing the quality of synthetic-aperture radar images. *Concurrency and Computation: Practice and Experience*, 33(13), p.e6239.
- [32] Dong, W., Zeng, H., Peng, Y., Gao, X. and Peng, A., 2022. A deep learning approach with data augmentation for median filtering forensics. *Multimedia Tools and Applications*, 81(8), pp.11087-11105.
- [33] Franchi, G., Fehri, A. and Yao, A., 2020. Deep morphological networks. *Pattern Recognition*, 102, p.107246.
- [34] Hu, Y., Belkhir, N., Angulo, J., Yao, A. and Franchi, G., 2022. Learning deep morphological networks with neural architecture search. *Pattern Recognition*, 131, p.108893.
- [35] yahia Ibrahim, N. and Talaat, A.S., 2022. An enhancement technique to diagnose colon and lung cancer by using double CLAHE and deep learning. *International Journal of Advanced Computer Science and Applications*, 13(8).
- [36] Islam, T., Hafiz, M.S., Jim, J.R., Kabir, M.M. and Mridha, M.F., 2024. A systematic review of deep learning data augmentation in medical imaging: Recent advances and future

research directions. *Healthcare Analytics*, p.100340.

- [37] Tammina, S., 2019. Transfer learning using vgg-16 with deep convolutional neural network for classifying images. *International Journal of Scientific and Research Publications (IJSRP)*, 9(10), pp.143-150.
- [38] Mascarenhas, S. and Agarwal, M., 2021, November. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification. In *2021 International conference on disruptive technologies for multi-disciplinary research and applications (CENTCON) (Vol. 1, pp. 96-99)*. IEEE.
- [39] Sinha, D. and El-Sharkawy, M., 2019, October. Thin mobilenet: An enhanced mobilenet architecture. In *2019 IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON) (pp. 0280-0285)*. IEEE.
- [40] Dong, K., Zhou, C., Ruan, Y. and Li, Y., 2020, December. MobileNetV2 model for image classification. In *2020 2nd International Conference on Information Technology and Computer Application (ITCA) (pp. 476-480)*. IEEE.
- [41] Wang, X., Li, J., Tao, J., Wu, L., Mou, C., Bai, W., Zheng, X., Zhu, Z. and Deng, Z., 2022. A recognition method of ancient architectures based on the improved inception v3 model. *Symmetry*, 14(12), p.2679.
- [42] Meyes, R., Lu, M., de Puiseau, C.W. and Meisen, T., 2019. Ablation studies in artificial neural networks. *arXiv preprint arXiv:1901.08644*.
- [43] Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L. and Ridella, S., 2012, April. The K'in K-fold Cross Validation. In *ESANN (Vol. 102, pp. 441-446)*.
- [44] Wong, T.T. and Yeh, P.Y., 2019. Reliable accuracy estimates from k-fold cross validation. *IEEE Transactions on Knowledge and Data Engineering*, 32(8), pp.1586-1594.