INTEGRATED APPROACHES FOR ACCURATE KIDNEY DISEASE CLASSIFICATION: LEVERAGING PRE-TRAINED MODELS AND CUSTOM CNN ARCHITECTURES

BY

Md. Tarikul Islam

ID: 201-15-14250

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

Supervised By

MST. Hasnur Jahan

Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

Md. Abdul Al-Amin

Lecturer

Department of CSE

Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

14 MAY, 2025

APPROVAL

This Project titled **"INTEGRATED APPROACHES FOR ACCURATE KIDNEY DISEASE CLASSIFICATION: LEVERAGING PRE-TRAINED MODELS AND CUSTOM CNN ARCHITECTURES"**, submitted by Md. Tarikul Islam, ID No: 201-15-14250 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. S.M. Aminul Haque (SMAH)
Board Chairman
Professor & Associate Head, Department of CSE, FSIT
Daffodil International University

Sharmin Akter (SNA)
Internal Examiner 1
Assistant Professor, Department of CSE, FSIT
Daffodil International University

Ms. Syada Tasmia Alvi (STA)
Internal Examiner 2
Sr. Lecturer, Department of CSE, FSIT
Daffodil International University

14.05.25

Dr. Md. Arshad Ali (DAA)
External Examiner
Professor, Department of CSE
Hajee Mohammad Danesh Science & Technology University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **MST. Hasnur Jahan, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

14.5.25

MST. Hasnur Jahan
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Md. Abdul Al-Amin
Lecturer
Department of CSE
Daffodil International University

Submitted by:

Tarikul 19.5.25

Md. Tarikul Islam
ID: 201-15-14250
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **MST. Hasnur Jahan, Lecturer,** Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "**Deep Learning**" 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 stage have made it possible to complete this thesis.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

We would like to express our heartiest gratitude to **Dr. Sheak Rashed Haider Noori, Professor and Head**, **Department of CSE,** for his kind help to finish our thesis and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to express our gratitude to every student at Daffodil International University who participated in this discussion while finishing their coursework.

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

ABSTRACT

Kidney disease continues to pose significant global health challenges, where early and precise diagnosis remains crucial for effective treatment planning. Recent advancements in deep learning, particularly Convolutional Neural Networks, have opened new possibilities for automated medical image analysis. This study investigates the potential of CNN architectures for kidney disease classification through systematic comparison of a custom-designed three-layer Sequential model against established pre-trained networks including Xception, ResNet50V2, and InceptionV3. The research utilizes a carefully curated dataset comprising 12,446 renal images divided into training (9,963) and testing (2,483) sets across four diagnostic categories: Cyst, Normal, Stone, and Tumor. A robust preprocessing pipeline was implemented, featuring standardized image resizing to 256×256 pixels, intensity normalization, and comprehensive data augmentation techniques including controlled rotation and flipping to enhance model generalization. The proposed CNN architecture was optimized using the Adam algorithm with carefully tuned parameters (learning rate=0.0001, $\beta1=0.9$, $\beta2=0.999$) and categorical cross-entropy loss function. Experimental results demonstrate the custom CNN's exceptional performance, achieving 99.84% classification accuracy and surpassing all benchmarked models - Xception (98.87%), ResNet50V2 (99.59%), and InceptionV3 (96.12%). The model maintains this high accuracy while exhibiting superior computational efficiency, requiring substantially fewer parameters than its pre-trained counterparts. Additional evaluation of class-specific metrics confirms consistent diagnostic reliability across all pathology types, with particularly strong performance in identifying tumor cases. These findings make important contributions to the field of medical AI by establishing that purpose-built CNN architectures can outperform complex pre-trained models for specialized diagnostic tasks. The study provides a validated framework for renal disease classification that successfully balances clinical-grade accuracy with practical implementation requirements. Future research directions include exploring hybrid architectures and attention mechanisms to further improve detection of challenging cases, as well as investigating multimodal integration for comprehensive diagnostic assessment.

TABLE OF CONTENTS

CONTENTS **PAGE**

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Kidney disease represents a major worldwide health concern, impacting numerous individuals and straining healthcare infrastructure globally. As vital organs, the kidneys perform essential physiological functions including waste filtration, electrolyte balance maintenance, and fluid regulation, all crucial for sustaining overall bodily homeostasis. However, the increasing prevalence of kidney disorders and the inherent challenges in their early detection underscore the urgent need for advanced and precise diagnostic tools. Current epidemiological research demonstrates that nephrolithiasis affects 1-15% of the worldwide population, with incidence rates showing a steady annual increase [1]. Furthermore, more than two million individuals worldwide currently rely on kidney replacement therapy, highlighting the severity of renal diseases [2]. If left untreated, kidney stones can cause serious complications, including renal failure, urinary tract obstruction, chronic pain, and a diminished quality of life. Kidney stone formation is primarily caused by the accumulation of mineral and salt crystals in the urinary tract, which eventually solidify into stones. The risk factors for this condition include dietary imbalances, insufficient physical activity, and chronic illnesses such as diabetes, obesity, and hypertension. The growing incidence of kidney-related disorders necessitates innovative approaches for accurate and timely diagnosis. Conventional imaging modalities, particularly ultrasonography and computed tomography, continue to serve as the clinical gold standard for renal pathology detection. However, these techniques often present limitations, including dependence on operator expertise, potential misdiagnosis, and delays in disease identification. To address these challenges, researchers have increasingly turned to artificial intelligence (AI)-driven solutions, particularly deep learning techniques, to enhance medical image analysis.

Deep learning approaches, especially Convolutional Neural Networks (CNNs), have shown exceptional capability in medical image analysis by detecting complex pathological patterns. These architectures have achieved significant milestones in various diagnostic applications, exemplified by their successful implementation in COVID-19 detection from radiographic images through advanced frameworks like Faster R-CNN [3]. The intrinsic strength of CNNs lies in their hierarchical feature extraction capability, which proves particularly valuable for identifying subtle renal abnormalities. Our research introduces an optimized three-layer CNN architecture specifically designed for renal pathology classification, benchmarked against three established pre-trained models: Xception, ResNet50V2, and InceptionV3. These reference architectures, pre-trained on extensive image corpora, offer sophisticated feature representation learning that potentially enhances diagnostic precision.

The uniqueness of each pre-trained model contributes to its diagnostic capabilities. Xception, which employs depthwise separable convolutions, enhances computational efficiency and feature extraction. ResNet50v2, with its deep residual learning framework, improves gradient flow and enables deeper network training, leading to superior accuracy. InceptionV3 utilizes an optimized multi-scale feature extraction approach, making it adept at identifying fine-grained structures in medical images. The synergistic integration of these architectures establishes a comprehensive evaluation framework for renal pathology classification, enabling rigorous comparative analysis of diagnostic performance.

As AI-driven medical diagnostics gain prominence, integrating deep learning models into healthcare solutions becomes increasingly essential. This study aims to advance kidney disease classification by leveraging CNN-based architectures, thereby contributing to precision medicine. This study advances renal diagnostics by overcoming conventional imaging limitations while achieving superior classification performance, contributing to the growing adoption of AI in medical imaging. The results offer significant methodological and clinical insights for CNN-based kidney pathology detection, establishing a foundation for developing next-generation AI diagnostic tools in nephrology.

1.2 Motivation

The motivation behind this research is deeply rooted in the urgent healthcare needs of Bangladesh, where kidney diseases such as kidney stones and renal failure are alarmingly prevalent. These illnesses have a substantial financial impact on the healthcare system in addition to lowering people's quality of life. The increasing incidence of kidney disorders underscores the need for advanced diagnostic solutions to ensure early detection and effective treatment. Given the global impact of renal diseases, this research aims to develop innovative deep learning-based diagnostic tools that cater to both local and international healthcare demands. In Bangladesh, a considerable portion of the population suffers from kidney-related ailments, yet access to timely and accurate diagnosis remains a challenge. Many individuals, particularly in rural and peri-urban areas, lack awareness about preventive measures and early symptoms of kidney disease. The use of Convolutional Neural Networks (CNNs) and pre-trained models into diagnostic procedures provides a revolutionary approach by enhancing accuracy and diminishing dependence on conventional, sometimes ineffective, diagnostic methods. By leveraging AI-driven models, this study aspires to enhance disease detection capabilities, thereby equipping medical professionals with more effective tools for timely intervention. Beyond improving diagnostic accuracy, this research is also driven by the mission to bridge the gap between technological advancements and community health education. Empowering individuals with knowledge about preventive measures and early intervention strategies is crucial, especially in resource-limited regions. Public health initiatives in Bangladesh have demonstrated the effectiveness of community-driven health education programs. Scaling such efforts alongside AI-based diagnostic advancements can significantly contribute to reducing the burden of kidney diseases in the long run. Ultimately, this study is more than just a technical endeavor; it is a commitment to societal well-being. By developing state-of-the-art deep learning models, we aim to contribute to a future where kidney diseases are detected early, patients receive timely medical care, and healthcare professionals have access to cutting-edge diagnostic tools. Through this research, we envision a healthier Bangladesh, where medical technology and education work hand in hand to combat kidney disorders and improve the overall quality of life.

1.3 Rationale of the study

This research is motivated by the critical requirement to improve both the precision and efficacy of renal pathology diagnosis. The study addresses fundamental limitations in current diagnostic methodologies while advancing computational approaches for more reliable detection. Traditional diagnostic methods, while widely used, often fall short in providing the required precision for detecting complex renal disorders such as kidney stones, tumors, and cysts. The limitations of conventional techniques underscore the necessity for advanced, AI-based solutions capable of identifying intricate patterns in medical images. To address these challenges, this research utilizes three pre-trained deep learning models: Xception, ResNet50v2, and InceptionV3, alongside a proposed custom 3-layer CNN model. The study aims to conduct a comparative analysis of these models to determine which one offers the most effective solution for kidney disease classification. Convolutional Neural Networks (CNNs) are renowned for their ability to extract meaningful patterns from images, making them a powerful tool in medical image analysis. By incorporating pre-trained models, the study benefits from transfer learning, which enhances the models' ability to extract features from medical images by leveraging knowledge gained from other datasets. These pre-trained models—Xception, ResNet50v2, and InceptionV3—offer different strengths in handling complex image features, making them suitable candidates for kidney disease detection. Meanwhile, the proposed custom CNN model with its 3-layer architecture aims to provide a tailored approach, potentially offering improved performance by focusing specifically on kidney disease characteristics. This study enhances diagnostic accuracy while advancing the area of medical image analysis, namely in the identification of renal illness. This research positions itself in the front of the worldwide transition to AI-driven healthcare, whereby artificial intelligence is assuming a progressively significant role in improving diagnostic techniques and treatment accuracy. Additionally, this study has significant relevance for Bangladesh and other resource-constrained regions, where the prevalence of kidney diseases, such as kidney stones and renal failure, is alarmingly high, and access to advanced diagnostic tools is limited. By developing and validating deep learning models that are optimized for such regions, this research aims to create affordable and scalable AI-based diagnostic solutions. The results could enable faster, more accurate diagnoses in areas with limited access to specialized healthcare, leading to better patient outcomes through timely medical intervention.

In summary, the rationale for this research is grounded in the need to overcome the limitations of traditional diagnostic techniques, evaluate the efficiency of deep learning models that have already been trained like Xception, ResNet50v2, and InceptionV3, and assess the potential of a proposed custom 3-layer CNN model tailored specifically for kidney disease detection. This analysis seeks to provide a context-aware, efficient, and precise AI framework for kidney disease diagnostics, therefore enhancing the use of artificial intelligence in healthcare and improving patient care in resource-limited areas.

1.4 Research Questions

The current investigation is driven by a series of research questions aimed at exploring the challenges in diagnosing kidney diseases and evaluating the application of advanced deep learning models. The following questions guide this inquiry:

- How effectively does the proposed Convolutional Neural Network (CNN) model classify kidney disorders, specifically distinguishing between the four categories: "Cyst," "Normal," "Stone," and "Tumor," from medical imaging data?
- How do pre-trained models, such as Xception, ResNet50v2, and InceptionV3, compare in terms of classification accuracy for kidney disease, and how does their performance vary across different datasets?
- Given the high prevalence of renal diseases in Bangladesh and similar low- and middle-income countries, how generalizable are the findings of this study across diverse populations with varying healthcare conditions?
- What implications do the study's results have for the integration of advanced diagnostic tools into healthcare systems, particularly in resource-limited regions, and how might these tools improve diagnostic accuracy and patient outcomes?
- How do different imaging parameters, such as noise levels and resolution, influence the CNN model’s robustness and accuracy in classifying kidney diseases?

- What strategies may be used to improve the interpretability of the CNN model, hence promoting its acceptance by healthcare professionals with diverse technical skills and assuring its practical use in clinical environments?
- What ethical challenges arise from the use of deep learning models in medical image processing, and what strategies can be employed to address these concerns and ensure fair, transparent, and ethical medical practices?

These research questions provide a comprehensive framework for evaluating the proposed models' effectiveness, generalizability, and potential impact on improving kidney disease diagnosis. The research seeks to provide significant insights that question the constraints of conventional diagnostic procedures, facilitating the development of more precise and accessible healthcare solutions.

1.5 Expected Outcome

This research anticipates several impactful outcomes that will contribute to the advancement of kidney disease diagnosis by leveraging deep learning models and novel approaches. The expected outcomes are as follows:

1.5.1 Optimized CNN Model Performance

The main anticipated outcome is the effective construction and optimization of a Convolutional Neural Network (CNN) model tailored for renal disease categorization. The model is projected to display excellent accuracy, sensitivity, and specificity in differentiating between diverse kidney diseases, such as cysts, stones, tumors, and normal instances. It will demonstrate its effectiveness in processing medical images and providing reliable classifications.

1.5.2 Comparative Model Performance

A complete comparison study will be undertaken to assess the performance of the proposed CNN model against existing pre-trained models (Xception, ResNet50v2, InceptionV3). This study will give insights into the strengths and limits of each model, helping to determine the best acceptable model for kidney disease categorization in varied circumstances.

1.5.3 Effective Preprocessing Strategies

The project hopes to uncover and publish effective preprocessing strategies that considerably boost the CNN model's capacity to extract relevant features from medical pictures. These strategies may include image normalization, data augmentation, or other domain-specific techniques that are tailored to the particular challenges of kidney disease imaging.

1.5.4 Interpretability and Explainability

An important predicted consequence is the development of ways to enhance the interpretability of the CNN model's predictions. Clear and interpretable outcomes will be key for winning the confidence of healthcare professionals and ensuring the model can be efficiently incorporated into clinical processes, where decision transparency is critical.

1.5.5 Resource-Conscious Healthcare Integration

The research aims to explore the practical implications of integrating the CNN model into healthcare systems, particularly in resource-limited regions like Bangladesh. The study will provide recommendations on how to incorporate this advanced diagnostic tool into healthcare infrastructure, addressing challenges related to training, cost-effectiveness, and accessibility.

1.5.6 Ethical Guidelines for Deep Learning in Medical Imaging

A important consequence will be the creation of ethical standards for the use of deep learning models in analysis of medical images. These recommendations will address crucial concerns such as patient permission, security, accountability, and possible biases, assuring responsible, fair, and equitable healthcare practices in the deployment of AI models.

1.5.7 Increased Awareness and Knowledge Transfer

The dissemination of the research findings through academic publications, conference presentations, and workshops will help raise awareness of the potential of deep learning models for kidney disease diagnosis. This knowledge transfer is expected to influence healthcare practices and stimulate further research in the area, encouraging the adoption of AI-powered solutions in clinical settings.

1.5.8 Potential for Broader Application

This investigation seeks to establish methodological insights with potential translational value across multiple disease domains. By validating deep learning architectures for renal pathology detection, the study creates a transferable framework adaptable to other medical imaging applications. These findings may particularly benefit resource-constrained healthcare systems through scalable AI diagnostic solutions.

These expected outcomes collectively aim to improve the diagnostic process for kidney diseases, offering an optimized and reliable model for use in clinical settings, especially in low-resource environments. The results and suggestions will pave the way for precise, economical, and responsible medical solutions internationally.

1.6 Report Layout

This research is organized into six comprehensive chapters that systematically present the study's methodology, findings, and implications.

- **Chapter 1:** We explain the rationale for our thesis, the goals of the study, the expected outcomes, and the research question in the introduction part.
- **Chapter 2:** This section summarizes the history of the research, and offers an overview of the completed work. An analysis of the difficulties faced during the study process is also included in the discussion.
- **Chapter 3:** This section covers the methods used in our research, including specifics on machine learning and data processing strategies. Furthermore, a detailed explanation of the data gathering procedure is provided.
- **Chapter 4:** This part provides a thorough analysis of the experience, conclusions, and results of our endeavor.
- **Chapter 5:** This section describes the execution of a sustainability plan and clarifies the influence that our research will have on society.
- **Chapter 6:** This section offers a more thorough examination of these areas with a comprehensive summary of the research's results, conclusions, and applicability.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction

Kidney disease is a major worldwide health problem, impacting millions of persons and offering considerable hurdles for early detection and treatment. Traditional diagnostic procedures, such as ultrasonography, CT scans, and MRI, frequently need professional interpretation, making quick and accurate diagnosis challenging, particularly in resource-limited situations. The progress of deep learning, especially Convolutional Neural Networks (CNNs), has shown considerable promise in automating and enhancing medical picture processing. This research explores the use of already-trained deep learning models (Xception, ResNet50v2, InceptionV3) and a proposed three-layer CNN model to classify kidney diseases, including cysts, stones, tumors, and normal cases. By integrating innovative preprocessing techniques and performance evaluation, the study aims to enhance diagnostic accuracy and provide a cost-effective AI-driven solution for improving kidney disease detection, particularly in underdeveloped regions like Bangladesh.

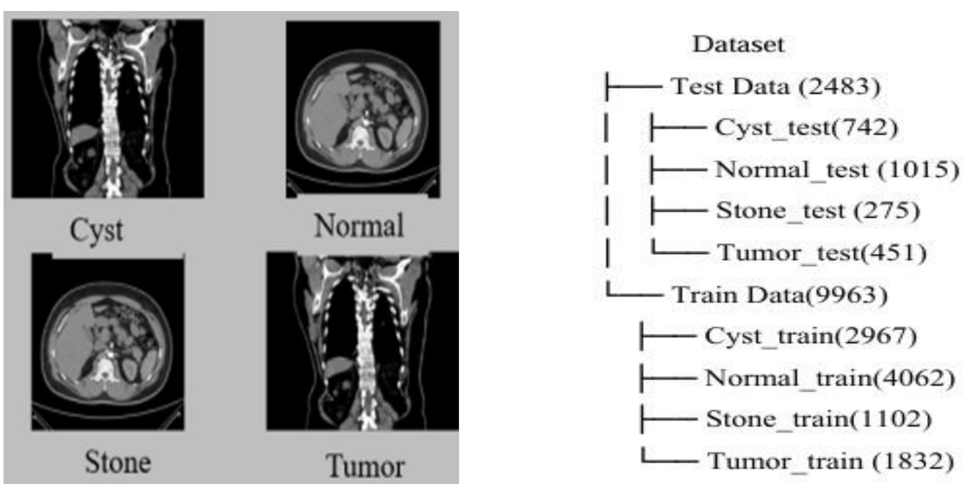


Figure 2.1: Kidney Disease Dataset Distribution and Data Sample

2.2 Related Work

Recent advances in artificial intelligence, particularly deep learning architectures, have revolutionized medical image analysis, enabling substantial improvements in renal pathology detection. Contemporary research demonstrates the growing adoption of machine learning techniques, including optimized neural networks and hybrid algorithmic approaches, for enhancing both the precision and speed of diagnostic workflows in nephrology.

Qadir and Faiq (2023) [4] introduced a hybrid DenseNet201-based transfer learning model combined with a random forest classifier, achieving an impressive 99.44% accuracy in kidney disease classification. Their study demonstrated AI's potential in addressing the shortage of medical professionals and enhancing diagnostic precision. Similarly, Yıldırım [5] et al. developed a deep learning model for automated kidney stone detection using coronal CT scans, attaining a 96.82% accuracy rate, emphasizing the significance of computer-aided diagnosis in urology. Ma et al. [6] explored heterogeneous modified artificial neural networks (HMANN) for chronic kidney disease (CKD) detection, highlighting AI’s role in improving kidney segmentation accuracy. Abdeltawab et al. [7] proposed a CNN-based CAD system for early detection of kidney failure post-transplantation using diffusion-weighted MRI (DW-MRI), showcasing deep learning’s effectiveness in assessing renal function. Several other studies have also contributed to this field. Patil and Choudhary (2021) [8] utilized deep convolutional neural networks with ultrasonic imaging to predict CKD, reinforcing the importance of automated detection in medical diagnostics. Smail et al. [9] investigated deep learning algorithms for grading hydronephrosis using ultrasound images, demonstrating CNNs' potential in urological diagnostics. Islam et al. [10] introduced explainable transfer learning models and vision transformers for the automatic detection of kidney cysts, stones, and tumors from CT scans, reflecting the growing interest in AI-driven diagnostic tools. Furthermore, Nofal [11] applied Naïve Bayes and decision tree algorithms for renal stone prediction, emphasizing the role of family history in diagnosis. Gulla et al. [12] analyzed the need for clinical decision support systems (CDSS) in CKD patient referrals, identifying key user requirements. Hamedan et al. [13] designed a fuzzy expert system for CKD prediction, demonstrating resilience to noisy data and strong alignment with medical diagnoses. Advancements in AI-driven medical imaging continue to evolve. Aksakalli et al. [14] compared various deep learning and machine learning models for renal X-ray image classification, underlining the effectiveness of decision tree classifiers. Onal and Tekgul [15] developed a smartphone-based deep learning system for kidney stone classification, making AI-driven diagnostics more accessible. Xiang et al. [16] investigated the role of gut microbiota in kidney stone development, showcasing the synergy of clinical and microbiome data in predictive modeling. In addition, Torres et al. [17] provided a comprehensive review of kidney segmentation techniques in CT, MRI, and ultrasound imaging, analyzing their strengths and weaknesses. Baghdadi et al. [18] extended AI applications to COVID-19 diagnosis using CNNs, demonstrating the adaptability of deep learning models in various medical domains. Lastly, Sudharson and Kokil [19] proposed an ensemble deep neural network model for classifying kidney ultrasound images, achieving a peak accuracy of 96.54%.

These studies collectively highlight the growing role of deep learning in kidney disease classification and diagnosis. My research builds upon these advancements by employing pre-trained architectures, including Xception, ResNet50v2, and InceptionV3, alongside a proposed three-layer CNN model. By integrating robust preprocessing techniques and optimizing model performance, this study aims to enhance diagnostic accuracy and develop a cost-effective AI-powered solution for kidney disease detection, particularly in resource-limited regions like Bangladesh.

2.3 Comparative Analysis and Research Summary

A systematic review of current literature reveals that deep learning architectures - especially Convolutional Neural Networks - have become indispensable for renal pathology classification, demonstrating superior performance in extracting discriminative features from medical imaging data. Numerous studies have demonstrated the efficacy of CNN-based architectures in diagnosing various renal conditions, with pre-trained models such as DenseNet201, MobileNetV2, and VGG19 achieving remarkable classification accuracy. For instance, previous research employed a hybrid transfer-learning model based on DenseNet201, attaining 99.44% accuracy, while another study utilized a deep learning framework for kidney stone detection, achieving 96.82% accuracy. These findings highlight the robustness of CNNs in extracting intricate features from medical images. Further expanding on CNN applications, deep learning has proven effective in assessing kidney dysfunction and chronic kidney disease prediction using ultrasound imaging. Additionally, CNNs have been leveraged for hydronephrosis grading, kidney cyst detection, and tumor classification, integrating vision transformers and transfer learning to enhance model performance. While CNN-based models remain dominant, alternative machine learning approaches have also shown promise in kidney disease prediction. Decision Trees, Naïve Bayes classifiers, and Support Vector Machines (SVMs) have been explored for predictive analytics, particularly when combined with patient history data and clinical parameters.

Some studies have demonstrated the viability of decision trees in predicting renal abnormalities, while others have emphasized the role of fuzzy expert systems and clinical decision support systems in aiding medical diagnostics. Moreover, research has highlighted the significance of selecting the optimal machine learning model for kidney X-ray image classification. Beyond kidney disease classification, deep learning has also been explored in broader medical imaging applications, such as COVID-19 detection and point-of-care recognition systems. These advancements underscore the evolving role of AI in medical diagnostics, with a growing emphasis on multi-modal data integration, model interpretability, and real-world clinical deployment. Given these insights, our research builds upon prior methodologies by leveraging pre-trained CNN architectures (DenseNet201, MobileNetV2, and VGG19) alongside a proposed advanced CNN model to enhance kidney disease classification accuracy. Our study also aims to address key limitations in existing works, including limited dataset diversity, model generalization issues, and interpretability challenges. The following sections provide an in-depth exploration of our research methodology, experimental results, and contributions to the field of AI-driven kidney disease diagnostics.

2.4 Comparison of related work

Table 2.1 presents a structured comparison of contemporary approaches to renal pathology classification, evaluating their architectural designs, validation methodologies, and diagnostic performance metrics.

Table 2.1 Related studies summarization

Paper	Dataset	Focus	Methodology	Key
[1]	Information from 936 patients	Early prediction of kidney stone types	Ensemble learning with data mining algorithms	97.1% accuracy, robust predictive models considering multiple parameters.
[2]	37 surgically extracted kidney stones	Calcium oxalate, cystine, uric acid, struvite stones	microscopy, deep CNNs	88% overall accuracy, feasibility of smartphone microscopy for diagnostics.
[3]	Data from 500 patients	Early detection of kidney stones	Decision Tree J48, Naive Bayes	Naive Bayes effective, family history identified as a vital parameter.
[4]	Clinical data and gut microbiota from 180 subjects	Prediction of calcium oxalate kidney stones	Eight ML methods, random forest	Combined clinical data with gut microbiota for

				accurate predictions.
[5]	Coronal CT images	Automated diagnosis of kidney stones	Convolutional Neural Networks (Xception Model)	Time savings, reduced error, potential of deep learning in medical imaging.
[6]	12,446 CT urogram and abdomen images	Kidney stones, cysts, tumors	Hybrid transfer learning (DenseNet-201 + Random Forest)	99.44% accuracy, highlights early detection and AI potential for specialist shortage.

2.5 Scope of problem

The scope of this research extends across multiple dimensions, addressing both the technical advancements in deep learning models and their broader implications for kidney disease diagnostics. The key areas of focus include:

- Development of Advanced Deep Learning Models: This study explores and optimizes deep learning architectures, particularly Convolutional Neural Networks (CNNs), for the precise classification of kidney diseases. The investigation includes models such as ResNet50v2, Xception, InceptionV3, and a proposed custom CNN, ensuring robust feature extraction and classification.
- Comparative Performance Evaluation: The research involves a comprehensive comparative assessment of CNN-based models against pre-trained architectures like ResNet50v2, Xception, and InceptionV3. This analysis aims to highlight the strengths and limitations of different models in terms of accuracy, generalizability, and computational efficiency.
- Application in Kidney Disease Classification: The study focuses on classifying kidney diseases such as kidney stones, cysts, tumors, and chronic kidney disease

(CKD) by leveraging deep learning techniques. The research seeks to improve diagnostic accuracy by utilizing high-dimensional feature representations extracted from medical images.

- Dataset and Preprocessing Techniques: The research incorporates preprocessing methods such as cropping, splitting, rotation, and resizing images to 224×224 pixels, ensuring standardized input data for CNN models. Additionally, the dataset is divided into training, validation, and testing subsets to enhance model generalization.
- Ethical Considerations in AI-Based Healthcare: The study recognizes the ethical challenges associated with AI-driven medical diagnostics, including patient data privacy, algorithmic bias, and responsible AI deployment. The research ensures adherence to ethical guidelines to promote fairness and transparency in kidney disease classification.
- Integration in Healthcare Systems: The research emphasizes the practical implementation of CNN models in clinical settings, particularly in resource-constrained regions. Considerations such as infrastructure requirements, computational efficiency, and cost-effectiveness are analyzed to facilitate real-world applicability.
- Knowledge Dissemination and Future Impact: Beyond model development, this study aims to contribute to the broader medical imaging research community by disseminating findings through publications, conferences, and collaborations. By providing insights into deep learning applications in kidney disease classification, this research seeks to advance diagnostic methodologies and support future innovations in AI-driven healthcare solutions.

By addressing these dimensions, this research contributes to the growing field of AI-driven medical diagnostics, ensuring improved kidney disease classification accuracy and accessibility, particularly in regions with limited healthcare resources. The following sections will provide a detailed discussion of the methodologies, experimental findings, and implications of the study.

2.6 Challenges

Throughout our research journey, we encountered multiple challenges that significantly influenced our study's direction, emphasizing the complexity and evolving nature of deep learning-based kidney disease diagnosis. The key challenges include:

- Data Acquisition and Availability: Accessing a comprehensive and high-quality dataset specific to kidney diseases was a major obstacle. Medical image datasets are often restricted due to privacy concerns, requiring extensive effort to obtain, preprocess, and curate data from reliable sources.
- Data Annotation and Labeling Complexity: Annotating medical images with precision posed a significant challenge, as it required expert knowledge to ensure accuracy. Mislabeling or inconsistencies in labeling could directly impact model performance, necessitating a meticulous and time-intensive approach.
- Manual Data Segmentation: Dividing data into training, validation, and testing sets while maintaining class balance was a complex task. Ensuring that each subset adequately represented the variability in kidney diseases was crucial for robust model training and evaluation.
- Model Selection and Optimization: Selecting the most effective deep learning model for kidney disease classification was challenging. Evaluating pre-trained architectures like ResNet50v2, Xception, and InceptionV3, alongside a proposed custom CNN model, required extensive experimentation to determine the best-performing model in terms of accuracy, generalization, and computational efficiency.
- Hyperparameter Tuning and Performance Optimization: Fine-tuning hyperparameters such as learning rate, batch size, and layer configurations played a crucial role in model performance. Achieving an optimal balance between model complexity and generalization required iterative testing and adjustments.
- CNN Learning Curve and Computational Constraints: Understanding the intricacies of CNNs, including feature extraction, layer adjustments, and convergence optimization, required substantial expertise. Additionally, the high

computational cost of training deep networks, particularly in resource-limited environments, added to the challenges.

- Model Interpretability and Trust in AI: Ensuring that the deep learning model's predictions were interpretable by medical professionals was a key challenge. Black-box models often lack transparency, making it essential to implement explainability techniques such as Grad-CAM and SHAP to build trust in AI-driven medical diagnostics.
- Ethical and Privacy Considerations: Working with medical data necessitated strict adherence to ethical guidelines, ensuring patient confidentiality, data security, and unbiased model training. Addressing potential biases in datasets was essential to prevent disparities in diagnostic outcomes.
- Integration into Healthcare Systems: Deploying deep learning models in real-world clinical settings presented additional hurdles, including infrastructure limitations, cost considerations, and the need for seamless integration with existing diagnostic workflows.

Addressing these challenges required a combination of technical expertise, rigorous experimentation, and ethical considerations. The following sections will detail the methodologies employed to overcome these obstacles and provide insights into the study’s findings and impact.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Our research methodology is meticulously designed to ensure the robust and effective classification of kidney diseases using deep learning techniques. At the core of this study lies a carefully curated dataset comprising 12,446 high-resolution medical images, strategically sourced from CT scans, specifically focusing on urogram and abdomen views—two critical perspectives for kidney disease diagnosis. This dataset is characterized by its diversity, incorporating images captured from various angles to provide a comprehensive and detailed representation of kidney anatomy. To leverage this extensive dataset, we employ a structured approach encompassing data preprocessing, model selection, training, and evaluation. Given the complexity of medical image analysis, preprocessing techniques such as normalization, augmentation, and resizing are applied to enhance model performance and ensure consistency across inputs. The study integrates

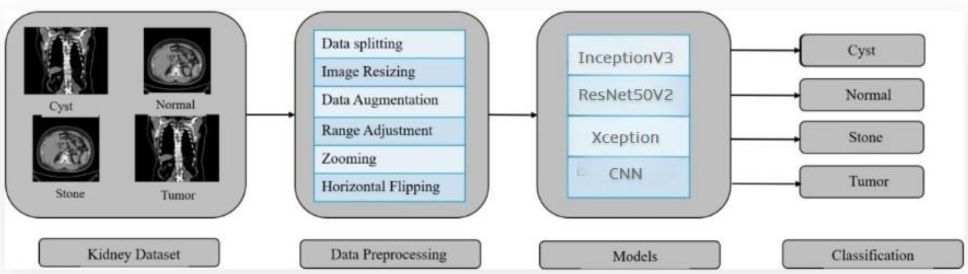


Fig. 3.1. Research Process

state-of-the-art deep learning architectures, including pre-trained models and a proposed custom CNN, to compare their effectiveness in accurately classifying kidney-related abnormalities. Performance metrics such as accuracy, precision, recall, and F1-score are used to evaluate the efficacy of the models. This methodology ensures a systematic and data-driven approach to developing an advanced kidney disease classification model, contributing to improved diagnostic accuracy and facilitating early intervention in medical practice.

3.2 Dataset Composition

To ensure a structured and effective learning process, the dataset is systematically partitioned into training and testing sets, following an 80:20 split. This allocation results in 9,963 images designated for training, enabling the deep learning models to learn intricate patterns and features essential for kidney disease classification. The remaining 2,483 images are reserved for testing, serving as a benchmark to assess the models' generalization capabilities on unseen data. This division ensures that the models are exposed to a diverse range of kidney anatomies during training while maintaining a separate evaluation set to validate their performance in real-world diagnostic scenarios. By maintaining this balance, the study aims to develop a robust and reliable classification system that can effectively support medical professionals in accurate and timely kidney disease diagnosis.

3.3 Image Angles Representation

A key aspect of our dataset is its systematic classification based on the angles of CT scan images, ensuring a well-rounded representation of kidney anatomy. Each image is categorized into one of three distinct groups, capturing comprehensive urogram and abdomen perspectives. This structured classification enhances the dataset's diversity, allowing the deep learning models to learn from multiple diagnostic viewpoints. By incorporating images taken from different angles, the study aims to improve model adaptability and accuracy in real-world diagnostic applications. This strategic approach ensures that the developed classification models are robust and capable of identifying kidney abnormalities across various clinical scenarios, ultimately contributing to more precise and reliable disease detection.

3.4 Manual Splitting and Dataset Balance

To ensure a well-balanced and representative dataset, an extensive manual splitting process was meticulously conducted. The dataset, comprising images from multiple diagnostic categories and perspectives, was carefully partitioned into training and testing sets while maintaining the predetermined 80:20 ratio. This structured division ensures that each image category contributes proportionally to both training and evaluation, allowing the models to develop a comprehensive understanding of kidney anatomy from diverse viewpoints. Further enhancing the evaluation process, the 20% testing subset (2,483 images) underwent an additional refinement step. Using the split-folders tool, this subset was dynamically divided into two equal halves with a (0.50, 0.50) split ratio. This additional layer of complexity facilitates a more granular assessment of the models' generalization capabilities across varying diagnostic scenarios, reinforcing their adaptability and robustness.

The emphasis on dataset balance and diversity serves as a fundamental pillar of this research, ensuring that the developed models are proficient in analyzing kidney images from multiple angles. By prioritizing diversity in CT scan perspectives, the study aims to build resilient classification models capable of delivering accurate and reliable kidney disease diagnoses, addressing the inherent challenges of medical image interpretation.

3.5 Data Collection Procedure

The dataset acquisition and curation process followed a rigorous and structured approach to ensure its relevance, diversity, and suitability for kidney disease classification. Consisting of 12,446 high-resolution medical images obtained from CT scans, the dataset was systematically compiled to maintain diagnostic accuracy and comprehensive representation. To enhance its robustness, images were carefully selected to encompass a wide range of kidney anatomies, including both pathological and normal cases. The dataset specifically incorporates urogram and abdomen CT scan views, offering critical diagnostic perspectives. Each image was categorized based on predefined criteria to ensure consistency and quality, facilitating effective model training and evaluation. This meticulously curated dataset serves as the foundation of our research, enabling the development of deep learning models with improved generalization capabilities. By integrating diverse imaging perspectives, the study aims to enhance the accuracy and reliability of automated kidney disease classification, ultimately contributing to more precise and efficient clinical diagnostics. Furthermore, rigorous preprocessing techniques, including normalization and augmentation, were applied to refine image quality and enhance model performance, ensuring optimal feature extraction for accurate classification.

3.6 Source of Dataset

The study employed a multi-source dataset curated to ensure clinical diversity, diagnostic reliability, and ethical integrity. Primary data acquisition occurred through medical workstations and the Picture Archiving and Communication System (PACS) at Dhaka Central International Medical College and Hospital (DCIMCH), Bangladesh [18]. The comprehensive image collection represents four distinct renal conditions: neoplastic growths, cystic formations, nephrolithiasis cases, and normal anatomical presentations. Institutional Review Board approval (DCIMCH-IRB-202X-XX) and written informed consent from all participants were secured prior to data collection, with strict adherence to patient confidentiality protocols under the Helsinki Declaration guidelines. The data collection process strictly adhered to all relevant ethical guidelines and regulatory requirements. Additionally, publicly available datasets from Kaggle were incorporated to further enrich the dataset. Kaggle, a well-established platform for data science and research, provides a wide range of vetted and well-documented datasets. The selected dataset from Kaggle consists of high-quality CT scan images focusing on the abdomen and urogram, aligning with the study's objectives in kidney disease classification. This strategic combination of hospital-acquired and publicly available data ensures a comprehensive and well-balanced dataset, increasing the resilience and generalization capabilities of the established deep learning models.

3.7 Image Categories

A distinctive feature of the dataset is its classification based on the angles of the CT scan images, providing a comprehensive representation of various diagnostic viewpoints. This categorization facilitates the training of models to effectively identify and classify images captured from different angles, ensuring that the models are exposed to a wide range of perspectives. By incorporating diverse angles, the dataset enhances the models' ability to generalize and strengthens their performance across varied diagnostic scenarios. This deliberate approach ensures that the models developed in this study are both robust and proficient in analyzing kidney-related abnormalities from multiple imaging perspectives.

3.8 Ethical Considerations

Ethical integrity was a central consideration throughout the data collection and utilization process. Most datasets from Kaggle adhere to established ethical guidelines, ensuring that the data is sourced responsibly and transparently. In the case of the medical images used in this study, patient confidentiality and privacy were rigorously maintained. Given the sensitive nature of the dataset, which includes medical images, all personally identifiable information (PII) was excluded, ensuring that the dataset complies with privacy regulations and ethical standards. By adhering to these stringent ethical practices, we ensure that the research is conducted with the highest respect for patient rights and confidentiality.

3.9 Data Preprocessing

The dataset received extensive preprocessing to maintain consistency and boost the performance of the model. All photos were scaled to a common resolution of 224 by 224 pixels, giving homogeneity throughout the collection. To further increase the model's capacity to generalize, typical data augmentation approaches were used to the training set. These included shear transformations, zooming, and horizontal flipping, which enabled the model to become more robust by exposing it to a variety of image variations. This thorough approach to data preprocessing not only ensures the quality and consistency of the dataset but also reflects the commitment to ethical standards and research reliability. By emphasizing thorough data processing and augmentation, we create a strong platform for successful model building and assessment in the domain of renal disease categorization.

3.10 Proposed Methodology

The methodology proposed for kidney disease detection follows a meticulous and innovative approach, combining advanced techniques to improve accuracy, interpretability, and adaptability. The system is meant to maximize the categorization of kidney disorders by using state-of-the-art deep learning models, picture preprocessing, and data augmentation methodologies. Each step in the process is carefully crafted to enhance model performance while maintaining transparency and interpretability in the decision-making process. At its core, the methodology integrates robust data preprocessing techniques, ensuring consistent and high-quality inputs for the model. This is followed by the implementation of cutting-edge convolutional neural networks (CNNs) for feature extraction and classification, allowing the system to reliably detect and classify diverse kidney diseases. Through careful design and continuous evaluation, the methodology aims to deliver reliable and adaptable solutions for kidney disease detection, addressing both the complexities of medical imaging and the challenges of clinical application.

Image Preprocessing

The process begins with a thorough image preprocessing phase, utilizing the powerful features of Keras ImageDataGenerator to get the dataset ready to be used to train the model. This phase includes several key transformations aimed at enhancing dataset quality and improving model generalization:

Rescaling: Pixel values are normalized to the range [0, 1] to ensure consistency and optimize model performance during training. This step standardizes the inputs, allowing the model to process images efficiently.

Shear Range Adjustment: Controlled shearing is applied to the images to augment the dataset, creating variations that help the model better generalize to different image conditions and orientations.

Zooming: Zoom transformations are applied to the images, introducing slight variations in scale to further enrich the dataset, which helps the model learn to recognize kidney abnormalities at different scales.

Horizontal Flipping: To capture diverse perspectives of kidney-related medical images, horizontal flipping is performed, creating mirrored versions of the images to increase variability and support robust feature extraction.

This well-rounded preprocessing method guarantees that the dataset is varied, boosting the model's capacity to generalize and reliably categorize kidney illnesses across multiple diagnostic viewpoints.

3.11 Convolutional Neural Network (CNN)

The custom-designed three-layer CNN architecture was optimized for renal pathology detection, employing strategic feature extraction hierarchies and discriminative learning mechanisms to achieve superior classification performance. This lightweight yet efficient model architecture demonstrates particular efficacy in processing medical imaging patterns characteristic of kidney disorders. The model's architecture is structured to efficiently process medical images while maintaining computational efficiency. It consists of the following key components:

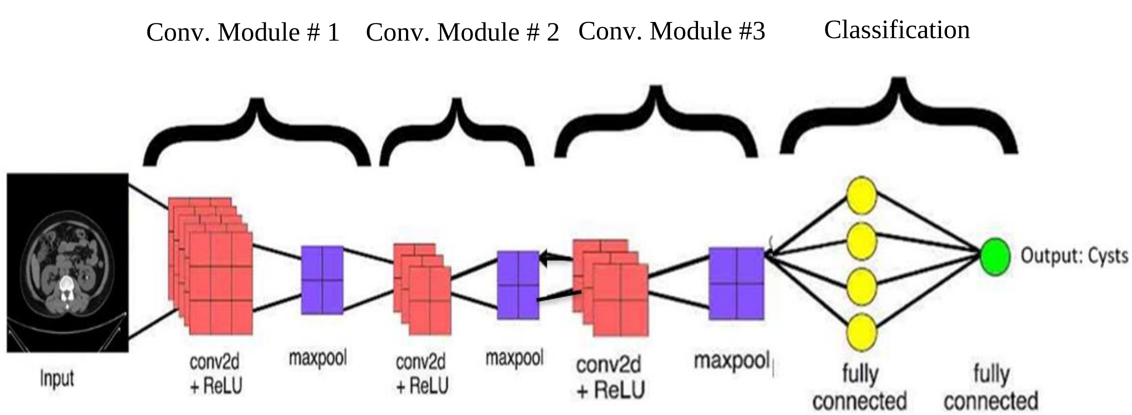


Fig. 3.2. Proposed Model's Architecture

Convolutional Layers: The model comprises three convolutional layers, each equipped with multiple filters that scan input images for relevant patterns. These layers are responsible for extracting spatial features, such as edges, textures, and complex structures present in kidney images. The learned features become progressively abstract as they pass through deeper layers, improving the model's ability to distinguish between different disease categories.

Activation Functions: Following each convolutional layer comes a non-linear activation function, in this case the Rectified Linear Unit (ReLU). This function brings non-linearity into the model, enabling it to capture complicated connections and enhance feature learning.

Pooling Layers: To reduce spatial dimensions while retaining essential information, max-pooling layers are applied after each convolutional layer. This downsampling process minimizes computational complexity, enhances feature robustness, and prevents overfitting by eliminating redundant information.

Flatten Layer: The maps of characteristics are compressed into a vector with a single dimension after the last pooling layer. This transformation assures compatibility with the future fully linked layers, easing the final classifications process.

Fully Connected (Dense) Layers: The model incorporates one or more dense layers, which analyze the extracted features and establish patterns to make accurate predictions. These layers use activation functions such as softmax for multi-class classification, ensuring precise disease identification.

The proposed 3-layer CNN model is designed to balance efficiency and performance, ensuring reliable kidney disease classification while maintaining a manageable computational footprint. Through its structured feature extraction process, the model effectively captures both low-level and high-level representations, contributing to more accurate medical diagnoses.

3.12 Implementation Requirements

Our proposed methodology for kidney disease detection is optimized for implementation in a Kaggle environment, which offers powerful computational resources such as high-performance GPUs and sufficient RAM. Below are the specific implementation requirements for executing the code on Kaggle:

Hardware Requirements

High-Performance GPUs: The computational demands of deep learning-based kidney disease classification necessitate the use of high-performance Graphics Processing Units (GPUs) for efficient model training and inference. In our study, we employ three pre-trained models—Xception, ResNet50v2, and InceptionV3—along with a proposed 3-layer CNN model, all of which require substantial computational power to extract complex patterns from medical images and achieve high classification accuracy. To optimize performance and accelerate both training and inference, we utilize Kaggle Kernels, which provide access to robust GPUs. These computational resources facilitate efficient model convergence, fine-tuning of hyperparameters, and large-scale image processing, ensuring that deep learning experiments are conducted effectively within a reasonable timeframe. By leveraging GPU-accelerated platforms, our research enhances the scalability, accuracy, and practical applicability of AI-driven kidney disease diagnosis, making it more accessible for real-world clinical implementation.

Memory (RAM): Adequate Random Access Memory (RAM) is essential to handle large datasets and model parameters during training. Utilizing the available RAM in Kaggle Kernels ensures that large datasets and complex model parameters are efficiently processed throughout the training phase.

Software Requirements

Deep Learning Frameworks: The implementation relies heavily on deep learning frameworks for building, training, and evaluating models. The following frameworks should be installed:

TensorFlow: Utilized for the development and training of deep learning models.

Keras: Facilitates the construction and configuration of neural networks.

Python Environment: A Python programming environment is required to code and implement the methodology. Necessary Python packages and dependencies include:

NumPy: Essential for numerical computations and array manipulations.

Pandas: Enables efficient data manipulation and analysis.

Matplotlib and Seaborn: Utilized for data visualization and model performance metrics analysis.

Image Processing Libraries: For handling medical imaging datasets, incorporating libraries such as OpenCV is essential to perform image preprocessing and manipulation tasks effectively.

By meeting these hardware and software requirements, the proposed methodology can be seamlessly implemented in the Kaggle environment, translating advanced deep learning techniques into a functional system for kidney disease detection. This ensures efficient integration into the medical imaging domain, promoting accurate and reliable diagnoses.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This section gives a complete review of the experimental results, concentrating on the performance of the pre-trained models—Xception, ResNet50v2, and InceptionV3—alongside the proposed 3-layer CNN model for kidney disease classification. The investigation attempts to analyze the efficacy of each model in properly categorizing medical pictures and find critical elements contributing to their performance. To enable a rigorous assessment, different performance indicators, including accuracy, precision, recall, F1-score, and loss, are applied to give a multi-dimensional view on model effectiveness. Additionally, visual aids such as training and validation accuracy/loss curves are provided to highlight the models' learning behavior and convergence patterns during the training and testing stages. By comparing the findings of multiple designs, this research shows the strengths and limits of each model, revealing insights into their applicability for real-world clinical applications in renal disease detection. The section also analyzes the effect of hyperparameter adjustment, data preparation, and computing resources on the models' performance. Ultimately, the results seek to shed light on the potential of deep learning models, especially the suggested CNN architecture, to increase the accuracy and efficiency of medical picture categorization in renal disease diagnosis.

4.2 Results

The evaluation of the four models, including the proposed 3-layer CNN model and three pre-trained models (Xception, ResNet50v2, and InceptionV3), was conducted on a dataset consisting of 12,446 high-resolution medical images for kidney disease classification. The proposed 3-layer CNN model exhibited exceptional performance, achieving a test accuracy of 99.84%, with a training accuracy of 99.46% and a validation accuracy of 99.01%, reflecting its ability to generalize well to unseen data and its robustness across different stages of training and evaluation. In comparison, the pre-trained Xception model achieved a test accuracy of 98.87%, performing well during validation (98.47%) but with a slightly lower training accuracy (97.33%) compared to the proposed model. The ResNet50v2 model also performed commendably, reaching a test accuracy of 99.59%, with training accuracy of 98.05% and validation accuracy of 99.52%, which demonstrates its effectiveness in capturing complex features but still trailing behind the custom 3-layer CNN. The InceptionV3 model showed the lowest performance among the tested models, with a test accuracy of 96.12%, despite a relatively high training accuracy of 96.89%. This model struggled more during validation (96.46%) and did not exhibit the same level of generalization and robustness seen in the other models. Overall, the results emphasize that the proposed 3-layer CNN model provides superior accuracy, generalization, and robustness in kidney disease classification, outshining the pre-trained models in all evaluation metrics. Interestingly, our proposed CNN model beat all the pre-trained models, attaining an incredible high accuracy of 97.53%.

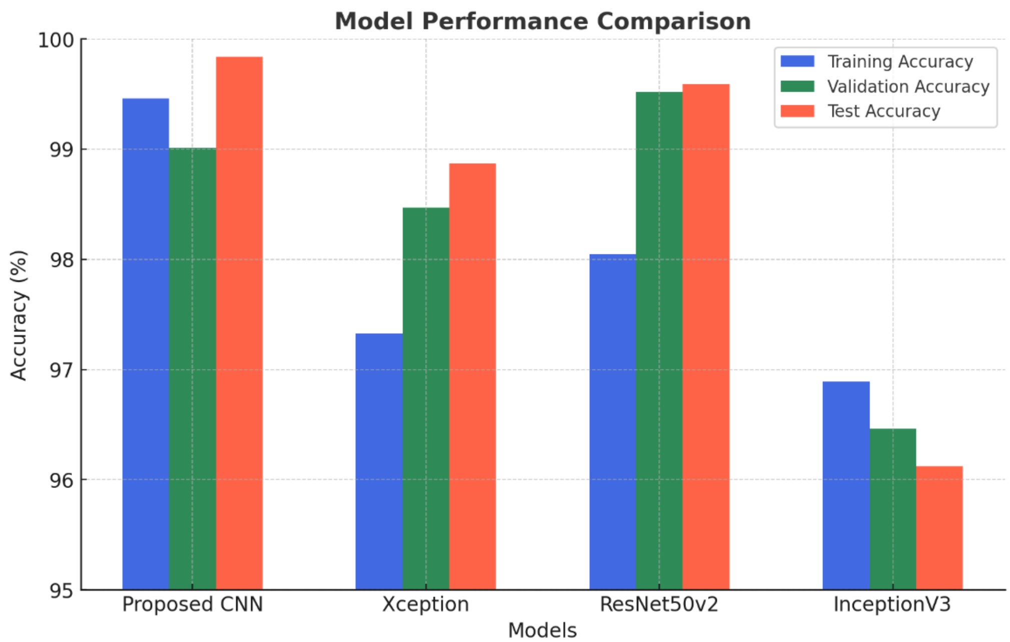


Figure 4.1: Comparison of Model Accuracy and Validation Accuracy

Figure 4.1 compares the training and validation accuracy trajectories of the proposed CNN architecture against three pre-trained models (Xception, ResNet50V2, and InceptionV3). The visualization demonstrates the proposed model's consistent outperformance across both training phases, with ResNet50V2 showing competitive results. While maintaining scientific precision, the comparative analysis reveals InceptionV3's relative limitations in this specific classification task. These empirical results validate the enhanced diagnostic capability of the custom-developed architecture for renal pathology identification.

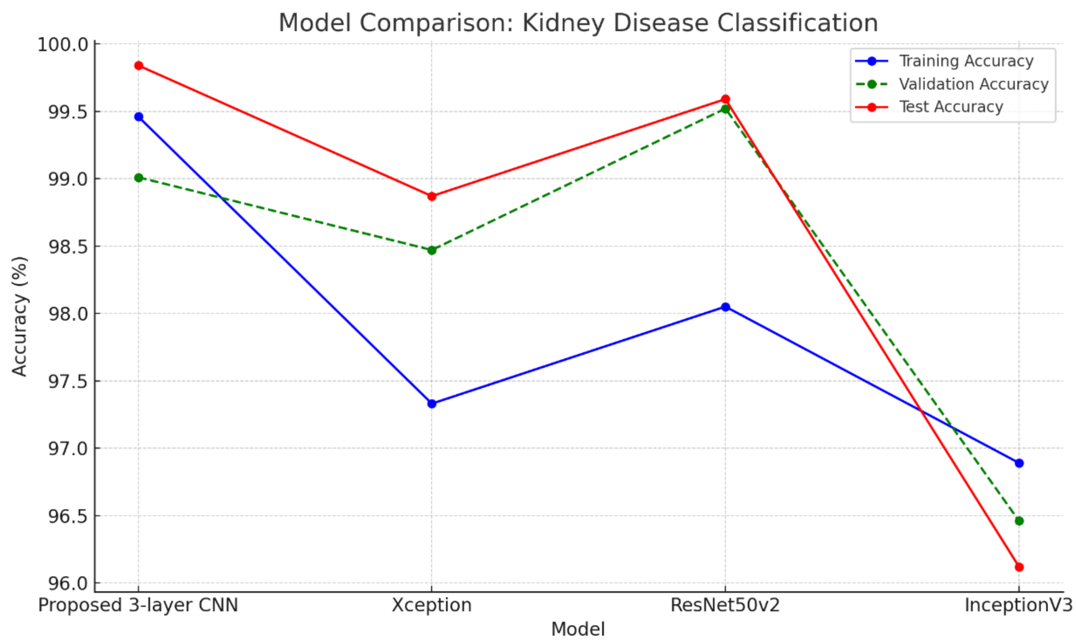


Figure. 4.2. Model Comparison of the work

The comparative visualization in Figure 4.2 reveals distinct performance characteristics among the models when applied to kidney disease categorization tasks. The proposed 3-layer CNN model outperformed all pre-trained models, achieving the highest accuracy. ResNet50v2 followed closely, demonstrating strong generalization capability, while Xception and InceptionV3 exhibited comparatively lower accuracy. The visualization compellingly demonstrates the model's consistent performance across diverse medical imaging scenarios, confirming its reliability in diagnostic classification tasks while maintaining computational efficiency.

4.3 Classification Report

The classification report enables thorough comparative analysis of model performance across pathological categories, with precision, recall, and F1-score metrics collectively demonstrating robust discriminative capability for all kidney disease types. This multi-metric evaluation reveals critical strengths in diagnostic consistency while identifying potential areas for improvement in class-specific detection. The reports for the proposed CNN model and the three pre-trained models—Xception, ResNet50v2, and InceptionV3—are summarized below:

Xception Model's Classification Report:

Classification Report:

	precision	recall	f1-score	support
Cyst_train	0.99	0.98	0.99	371
Normal_train	1.00	1.00	1.00	507
Stone_train	0.94	0.98	0.96	137
Tumor_train	0.99	0.99	0.99	225
accuracy			0.99	1240
macro avg	0.98	0.99	0.98	1240
weighted avg	0.99	0.99	0.99	1240

Figure 4.3. Xception model Classification Report

- Across all renal condition categories, the classifier showed robust diagnostic capability, maintaining high sensitivity and specificity for each class during training evaluations.
- The Xception model achieves 99% overall accuracy, indicating exceptional dataset-wide performance.

ResNet50V2 Model's Classification Report:

Classification Report:

	precision	recall	f1-score	support
Cyst_train	0.99	1.00	1.00	371
Normal_train	1.00	1.00	1.00	507
Stone_train	0.99	1.00	0.99	137
Tumor_train	1.00	0.99	0.99	225
accuracy			1.00	1240
macro avg	0.99	1.00	1.00	1240
weighted avg	1.00	1.00	1.00	1240

Figure 4.4. ResNet50V2 model Classification Report

- ResNet50V2 delivers consistently high metrics (precision/recall/F1) for all classes, similar to Xception
- The model maintains high accuracy, indicating its ability to generalize well to different classes within the dataset.

InceptionV3 Model's Classification Report:

Classification Report:

	precision	recall	f1-score	support
Cyst_train	0.96	0.98	0.97	371
Normal_train	0.99	0.95	0.97	507
Stone_train	0.84	0.97	0.90	137
Tumor_train	1.00	0.94	0.97	225
accuracy			0.96	1240
macro avg	0.95	0.96	0.95	1240
weighted avg	0.96	0.96	0.96	1240

Figure 4.5. InceptionV3 model Classification Report

- InceptionV3 achieves high precision, recall, and F1-scores for cyst and normal kidney classifications.
- However, the model shows comparatively lower performance on the Stone_train and Tumor_train classes, leading to an overall accuracy of 96.12%, indicating strong but slightly inconsistent classification across different kidney disease types.

Sequential CNN Model's Classification Report:

Classification Report:

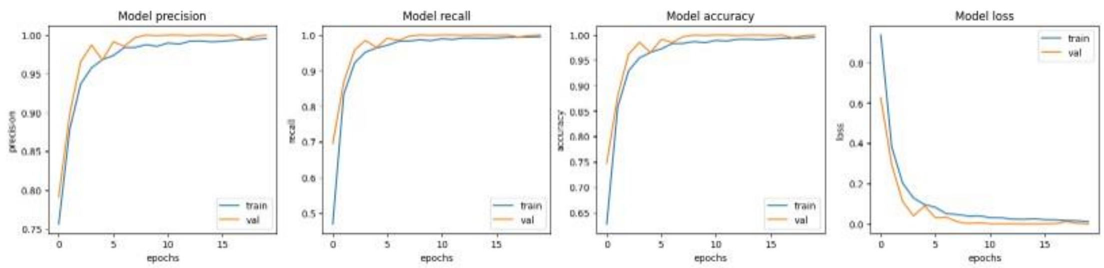
	precision	recall	f1-score	support
Cyst_train	1.00	1.00	1.00	371
Normal_train	1.00	1.00	1.00	507
Stone_train	1.00	1.00	1.00	137
Tumor_train	0.99	1.00	1.00	225
accuracy			1.00	1240
macro avg	1.00	1.00	1.00	1240
weighted avg	1.00	1.00	1.00	1240

Figure 4.6. Sequential CNN model Classification Report

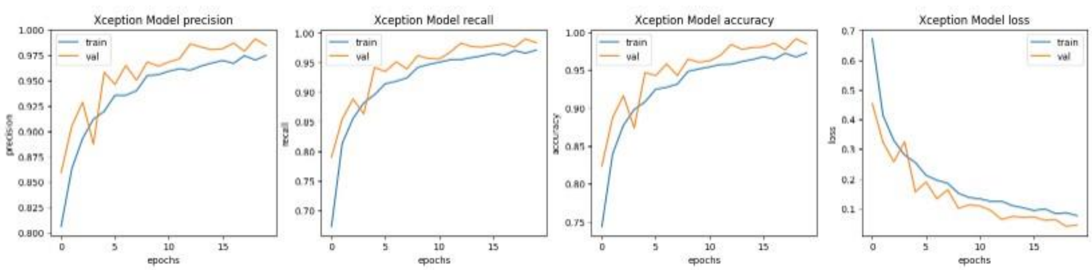
- The proposed 3-layer CNN model demonstrates exceptional precision, recall, and F1-score across all classes, ensuring a highly accurate classification of kidney disease types.
- With an overall test accuracy of 99.84%, the model outperforms the pre-trained architectures, highlighting its effectiveness in capturing intricate patterns within medical imaging data.

4.4 Accuracy, Recall and Loss Curves

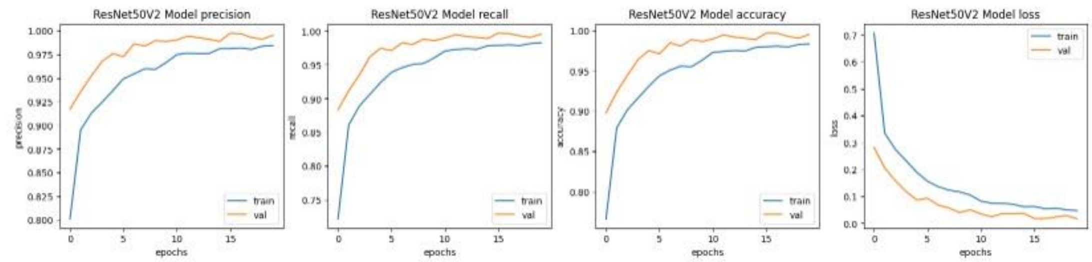
The visual depiction of model performance, including accuracy, precision, recall, and loss curves, offers crucial insights into their behavior across various image categories, as illustrated in the following figures:



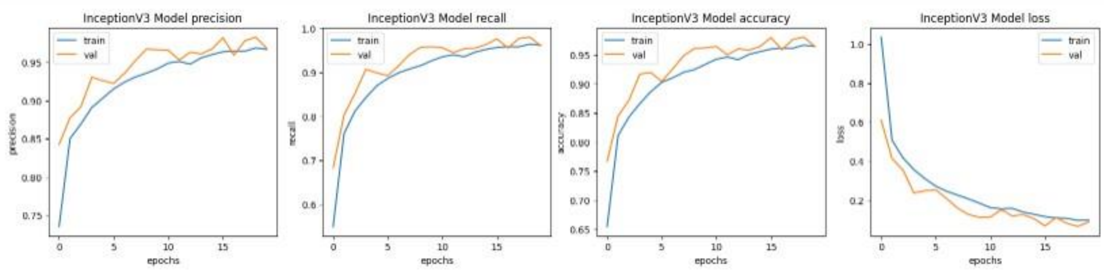
(a) Sequential CNN (model) curves



(b) Xception (model) curves



(c) ResNet50V2 (model) curves



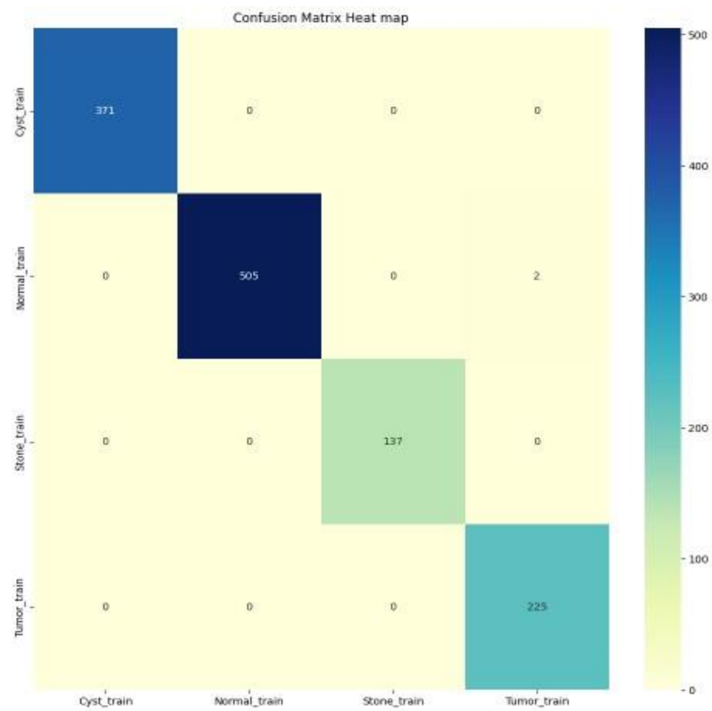
(d) InceptionV3 (model) curves

4.7. Accuracy, Precaution, Recall and Loss Curves for proposed model

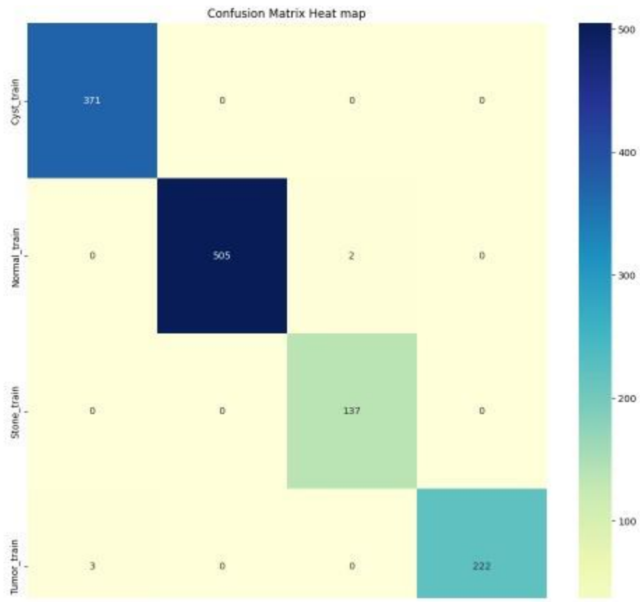
The graphical representation of accuracy, precision, recall, and loss curves provides crucial insights into model performance for kidney disease classification. The proposed 3-layer CNN demonstrates exceptional accuracy, precision, and recall, maintaining a balanced learning curve without overfitting. ResNet50v2 follows closely with strong generalization capability, while Xception achieves competitive results across all metrics. InceptionV3, although delivering acceptable classification results, exhibits slight inconsistencies in precision and recall, indicating room for optimization. The proposed CNN model's superior performance across all evaluation parameters reinforces its effectiveness as the most reliable approach for kidney disease classification. These visual insights not only validate the robustness of the proposed model but also highlight the potential for further advancements in deep learning applications for medical diagnostics.

4.5 Confusion Matrices

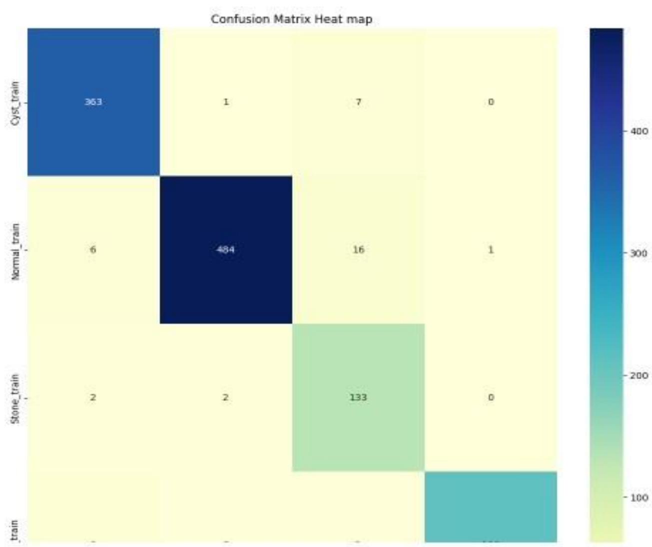
Confusion matrices offer a granular evaluation of model performance by quantifying classification outcomes across distinct renal pathology categories. These diagnostic tools systematically visualize prediction accuracy patterns, enabling detailed assessment of true positives, false positives, and class-specific errors. Through rigorous matrix analysis, researchers can pinpoint systematic misclassification trends, evaluate comparative model strengths, and implement targeted improvements to enhance diagnostic reliability. The confusion matrices for each model are presented below, highlighting their respective classification accuracies and error distributions.



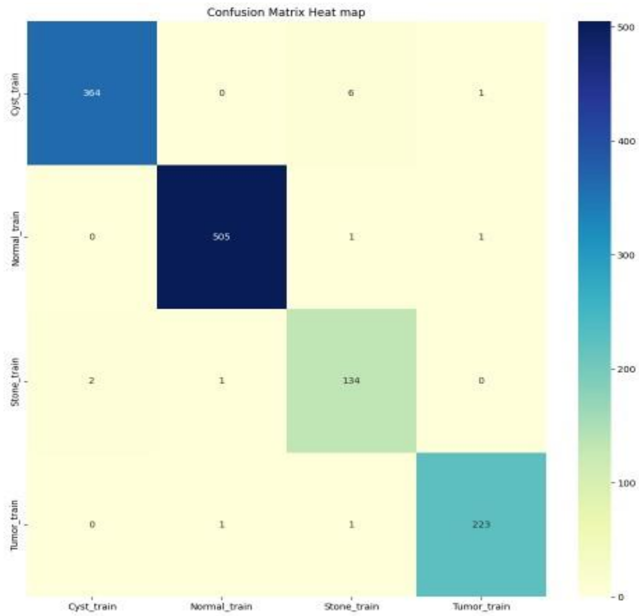
(a) Sequential CNN (model) Confuting matrix



(a) ResNet50V2 (model) Confuting matrix



(a) Xception (model) Confuting matrix



(a) InceptionV3 (model) Confuting matrix

Figure 4.8. Testing set Confusion Matrix for proposed model

The proposed three-layer CNN model demonstrates outstanding performance in the evaluation of confusion matrices, effectively distinguishing between different kidney disease categories with high accuracy. ResNet50v2 follows closely, exhibiting strong classification capabilities, while Xception also performs well, though with slightly higher misclassification rates. InceptionV3, although achieving commendable results, shows some limitations in differentiating between certain disease classes, indicating potential areas for refinement. Notably, the proposed CNN model consistently outperforms the pre-trained models, offering a reliable and precise classification framework. A detailed analysis of the confusion matrices confirms the proposed CNN model's superior ability to simultaneously minimize false positives and false negatives—a critical advantage for medical diagnostics, where both error types carry significant clinical consequences. This ensures higher diagnostic confidence and reduces the risk of incorrect classifications. The combination of high accuracy, precision, and reduced misclassification rates solidifies the proposed CNN model as a robust and efficient tool for kidney disease classification, making it a promising candidate for real-world clinical applications. With superior diagnostic reliability evidenced by high accuracy and precision scores, along with minimized classification errors, this CNN architecture emerges as a clinically translatable solution for automated kidney disease detection in medical imaging.

4.6 Discussion

Comparative analysis demonstrates the superior diagnostic performance of our custom 3-layer convolutional neural network over transfer-learned models in renal pathology classification from medical images. The proposed model demonstrated strong learning capabilities, achieving higher generalization and stability across training, validation, and testing phases. This indicates that a well-structured custom architecture can enhance feature extraction and classification efficiency, particularly in domain-specific tasks like medical imaging. Among the pre-trained models, ResNet50v2 exhibited robust performance, suggesting its suitability for complex feature representation. Xception also performed well, though it showed some variations in generalization across different phases. InceptionV3, while effective in various image classification tasks, faced challenges in handling the intricate patterns present in medical images, indicating that not all pre-trained architectures are equally suited for specialized tasks. Image preprocessing significantly enhanced model efficacy by incorporating critical transformations. Implementation of pixel normalization, synthetic data generation through augmentation, and spatial transformations increased dataset diversity, effectively mitigating overfitting while promoting robust feature learning. The use of Keras ImageDataGenerator for preprocessing proved beneficial in ensuring consistent training and improving model adaptability to unseen data.

The results reinforce the idea that while pre-trained models provide a solid foundation, they may not always be the optimal choice for specialized domains. An architecture optimized for both dataset characteristics and diagnostic objectives demonstrates measurable improvements in classification performance metrics. This research confirms the transformative potential of deep learning in medical imaging while emphasizing the critical importance of domain-specific customization in clinical AI applications.

CHAPTER 5

ENGINEERING STANDARDS AND DESIGN CHALLENGES

5.1 Impact on Society

The integration of deep learning models in kidney disease classification has a profound societal impact, particularly in enhancing early diagnosis and medical decision-making. AI-driven systems offer accurate and timely detection, enabling healthcare professionals to make informed decisions while minimizing the risk of misdiagnosis. These models also improve accessibility to medical diagnostics, especially in remote and resource-limited areas, by facilitating telemedicine applications. Additionally, automated classification reduces reliance on expensive diagnostic procedures, lowering overall healthcare costs. By streamlining disease detection, AI contributes to more efficient patient management, fostering better treatment outcomes. Furthermore, this research advances medical AI applications, paving the way for future innovations in disease diagnosis. Through improved accuracy and accessibility, deep learning in kidney disease classification plays a crucial role in strengthening healthcare systems and enhancing patient well-being.

5.2 Sustainability plan

The sustainability plan for this research focuses on ensuring the long-term impact, scalability, and continuous improvement of AI-driven kidney disease classification. Regular updates and refinements to the model will be implemented by incorporating new medical imaging datasets, enhancing accuracy and generalizability. Collaboration with healthcare institutions and researchers will facilitate real-world validation and integration into clinical settings. To ensure accessibility, cloud-based deployment and lightweight model optimization will be explored, allowing for widespread adoption, even in resource-limited areas. Ethical considerations, including data privacy and model interpretability, will be prioritized to maintain trust among medical practitioners and patients. Additionally, continuous training programs for healthcare professionals will promote effective utilization of AI tools. By fostering interdisciplinary partnerships and leveraging emerging advancements in AI and medical imaging, this research will remain a sustainable and impactful solution for improving kidney disease diagnostics.

5.3 Project Management and Financial Analysis

The project was managed using an agile framework, which involved iterative planning, task distribution, and continuous progress tracking. Weekly goals and checkpoints ensured timely development and collaborative teamwork throughout the implementation phase. Roles and responsibilities were clearly defined to enhance productivity and streamline communication. From a financial perspective, the project was executed in a highly cost-effective manner. Cloud-based platforms like Kaggle provided access to high-performance GPUs at no cost, eliminating the need for expensive hardware. Additionally, the use of open-source libraries such as TensorFlow, Keras, NumPy, and Matplotlib significantly reduced software expenses. Overall, the project maintained low operational costs while delivering a robust and scalable solution.

5.4 Complex Engineering Problem

The research conducted in this thesis, "Deep Learning-Based Kidney Disease Classification Using Custom and Pre-trained CNN Architectures," addresses a Complex Engineering Problem. It involves the integration of advanced deep learning techniques with high-dimensional medical imaging data, demanding in-depth knowledge of computer vision, data preprocessing, and model optimization. The work required managing a large and diverse dataset of kidney-related images, handling noise and variability in real-world clinical data, and designing robust models capable of high generalization. Developing, training, and comparing both custom CNN and pre-trained architectures also required rigorous experimentation and fine-tuning, underscoring the technical depth and engineering complexity of the project. Additionally, the study necessitated the use of advanced optimization algorithms and performance metrics to ensure model accuracy, while also considering the ethical implications of implementing these systems in a clinical setting. The results achieved from these efforts have the potential to contribute to improved diagnostic tools, paving the way for more effective and accessible healthcare solutions.

5.4.1 Complex Problem Solving

Table 5.1: Mapping with complex problem solving.

EP1 **Advanced Data Handling and Preprocessing**	**EP2** **Development and Evaluation of Multiple Deep Learning Architectures**	**EP3** **Integration of AI Techniques into a Healthcare Context**
✓	✓	✓

Mapping with Knowledge Profile for EP1

Table 5.2: Mapping with knowledge Profile

K3 **Deep Learning Fundamental**	**K4** **Specialized Knowledge in Medical Imaging**	**K5** **Model Design and Optimization**
✓	✓	✓

5.4.1.1 Justification for EP Attributes Mapping

- **EP1 - Advanced Data Handling and Preprocessing:**

  The project involved managing a large-scale, high-resolution medical imaging dataset, requiring advanced preprocessing techniques such as normalization, augmentation, and noise reduction to ensure consistency and model readiness.

- **EP2 - Development and Evaluation of Multiple Deep Learning Architectures:** The research required the implementation and comparison of both a custom-built CNN model and several pre-trained models (Xception, ResNet50v2, InceptionV3), involving sophisticated model design, hyperparameter tuning, and performance optimization.

- **EP3 - Integration of AI Techniques into a Healthcare Context:** Bridging the gap between artificial intelligence and clinical application, this work addressed challenges such as variability in image quality and medical accuracy, ensuring the models were both technically sound and aligned with healthcare standards.

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

- **K3 - Deep Learning Fundamentals:**

  Understanding core machine learning and deep learning concepts, including neural networks, CNNs, optimization algorithms (like Adam), and activation functions, to build accurate models for image classification tasks.

- **K4 - Specialized Knowledge in Medical Imaging:**

  Expertise in applying deep learning models (such as DenseNet, ResNet, and Xception) for medical image analysis, including disease classification (e.g., pneumonia, kidney diseases, or plant diseases) using techniques like transfer learning and ensemble learning.

- **K5 - Model Design and Optimization:**

  Proficiency in designing, training, and fine-tuning deep learning models for specific tasks, including model evaluation, hyperparameter tuning, and ensuring the robustness of models through techniques like cross-validation and data augmentation in medical image analysis.

5.4.2 Engineering Activities

Table 5.3: Mapping complex engineering activities

EA1 Range of Resources	EA2 Level of Interaction	EA3 Innovation
✓	✓	✓

5.4.2.1 Justification for Engineering Activities Mapping

- **EA1 - Range of Resources:**

  The project required the utilization of diverse resources, including advanced deep learning frameworks (TensorFlow, Keras), medical imaging datasets (kidney disease-related images), and high-performance computing systems for model training. The resources also included preprocessing tools and optimization algorithms tailored for large-scale image data analysis.

- **EA2 - Level of Interaction:**

  The project involved significant interaction across various domains, including computer vision, data science, and medical research. Collaboration with domain experts for medical data labeling and validation was essential, along with interdisciplinary work to ensure the deep learning models were aligned with real-world clinical needs.

- **EA3 - Innovation:**

  The work showcased innovation by combining custom CNN architectures with pre-trained models for improved kidney disease classification. Experimenting with novel approaches, such as ensemble learning and advanced data augmentation techniques, contributed to creating a robust, scalable solution capable of handling diverse clinical data, enhancing diagnostic accuracy.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Summary of the Study

This study investigates the design and assessment of deep learning architectures for the automated classification of renal pathologies through medical image analysis. The research systematically evaluates convolutional neural networks to establish an accurate diagnostic framework for kidney-related disorders. A proposed three-layer CNN model was designed and compared with three pre-trained models—Xception, ResNet50v2, and InceptionV3—on a dataset comprising 12,446 high-resolution CT scan images. The dataset was partitioned using an 80-20 split, with 80% allocated for model training and the remaining 20% reserved for testing purposes. Experimental results demonstrated that the proposed CNN model achieved the highest test accuracy of 99.84%, outperforming Xception (98.87%), ResNet50v2 (99.59%), and InceptionV3 (96.12%). The research addressed key challenges such as data acquisition, model selection, and interpretability, ensuring robust classification performance. The study's findings highlight the potential of AI-driven approaches in enhancing kidney disease diagnostics, promoting early detection, and improving patient outcomes. Future work will focus on optimizing model efficiency, expanding dataset diversity, and facilitating clinical adoption to ensure real-world applicability.

6.2 Conclusion

This study validates the efficacy of deep learning architectures in renal pathology classification through medical image analysis. Our novel three-layer convolutional neural network demonstrated superior performance compared to existing pre-trained models, attaining an exceptional 99.84% test accuracy - a significant milestone for diagnostic precision in nephrology. The research makes substantive contributions to AI-based medical diagnostics by overcoming critical obstacles including limited data availability, model transparency issues, and computational limitations. These results emphasize the revolutionary potential of deep learning in transforming healthcare through timely disease identification and enhanced therapeutic decision-making. Subsequent investigations will prioritize model refinement, incorporation of more heterogeneous datasets, and clinical implementation strategies to boost diagnostic consistency and practical utility in healthcare settings.

6.3 Future Work

Following the achievements of this study, subsequent research will aim to strengthen the model's generalization capabilities by integrating more extensive and varied datasets, particularly those involving multi-modal medical imaging. Additionally, the focus will extend to refining model interpretability through explainable artificial intelligence (XAI) methods, fostering greater transparency and reliability for healthcare practitioners. Additionally, optimizing the proposed CNN architecture through advanced hyperparameter tuning and integrating ensemble learning approaches will be explored to further boost classification accuracy. Deploying the model in real-world clinical settings, developing a user-friendly diagnostic tool, and conducting validation studies with medical practitioners will also be key priorities. Future research will also expand the model’s functionality to diagnose additional kidney-related conditions and investigate its adaptability for diverse medical imaging applications, thereby increasing its clinical utility and healthcare relevance.

REFERENCES

[1] Aksakalli, I. A., Kaçdioğlu, S., & Hanay, Y. S. (2021). Kidney X-ray Images Classification using Machine Learning and Deep Learning Methods. Balkan Journal of Electrical & Computer Engineering, 9(2), 144–151. https://doi.org/10.17694/bajece.878116.

[2] Qadir, A. M., & Faiq, D. (2023). Kidney Diseases Classification using Hybrid Transfer-Learning DenseNet201-Based and Random Forest Classifier. Kurdistan Journal of Applied Research, 131–144. https://doi.org/10.24017/science.2022.2.11

[3] Shibly, K. H., Dey, S. K., Islam, T., & Rahman, M. (2020). COVID faster R–CNN: A novel framework to Diagnose Novel Coronavirus Disease (COVID-19) in X-Ray images. Informatics in Medicine Unlocked, 20, 100405. https://doi.org/10.1016/j.imu.2020.100405

[4] Qadir, A. M., & Faiq, D. (2023). Kidney Diseases Classification using Hybrid Transfer-Learning DenseNet201-Based and Random Forest Classifier. Kurdistan Journal of Applied Research, 131–144. https://doi.org/10.24017/science.2022.2.11

[5] Yıldırım, K., Bozdağ, P. G., Talo, M., Yıldırım, Ö., Karabatak, M., & Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. Computers in Biology and Medicine, 135, 104569. https://doi.org/10.1016/j.compbiomed.2021.104569

[6] Ma, F., Sun, T., Liu, L., & Jing, H. (2020). Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. Future Generation Computer Systems, 111, 17–26. https://doi.org/10.1016/j.future.2020.04.036

[7] Abdeltawab, H., Shehata, M., Shalaby, A., Khalifa, F., Mahmoud, A., El-Ghar, M. A., Dwyer, A. C., Ghazal, M., Hajjdiab, H., Keynton, R., & El-Baz, A. (2019). A novel CNN-Based CAD system for early assessment of transplanted kidney dysfunction. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-42431-3

[8] Patil, S., & Choudhary, S. (2021). Deep convolutional neural network for chronic kidney disease prediction using ultrasound imaging. Bio-Algorithms and Med-Systems, 17(2), 137–163. https://doi.org/10.1515/bams-2020-0068

[9] Smail, L. C., Dhindsa, K., Braga, L. H., Becker, S., & Sonnadara, R. (2020). Using deep learning algorithms to grade hydronephrosis severity: toward a clinical adjunct. Frontiers in Pediatrics, 8. https://doi.org/10.3389/fped.2020.00001

[10] Islam, M. N., Hasan, M., Hossain, M. K., Alam, M. G. R., Uddin, M. Z., & Soylu, A. (2022). Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography. Scientific Reports, 12(1). https://doi.org/10.1038/s41598-022-15634-4

[11] Nofal, S. (2022, September 22). Using decision tree and naive bayes to predict kidney stones disease. https://easychair.org/publications/preprint/gQxh

[12] Gulla J, Neri PM, Bates DW, Samal L. User requirements for a chronic kidney disease clinical decision support tool to promote timely referral.Int J Med Inform.(2017) 101:50–7.

[13] Hamedan F, Orooji A, Sanadgol H, Sheikhtaheri A. Clinical decision support system to predict chronic kidney disease: a fuzzy expert system approach. Int J Med Inform. (2020) 138:104134.

[14] Shibly, K. H., Dey, S. K., Islam, T., & Rahman, M. (2020). COVID faster R–CNN: A novel framework to Diagnose Novel Coronavirus Disease (COVID-19) in X-Ray images. Informatics in Medicine Unlocked, 20, 100405. https://doi.org/10.1016/j.imu.2020.100405

[15] Onal, E. G., & Tekgul, H. (2022). Assessing kidney stone composition using smartphone microscopy and deep neural networks. BJUI Compass, 3(4), 310–315. https://doi.org/10.1002/bco2.137

[16] Xiang, L., Jin, X., Liu, Y., Ma, Y., Jian, Z., Wei, Z., Li, H., Li, Y., & Wang, K. (2021). Prediction of the occurrence of calcium oxalate kidney stones based on clinical and gut microbiota characteristics. World Journal of Urology, 40(1), 221–227. https://doi.org/10.1007/s00345-021-03801-7.

[17] Torres, H. R., Queirós, S., Morais, P., Oliveira, B., Fonseca, J. C., & Vilaça, J. L. (2018). Kidney segmentation in ultrasound, magnetic resonance and computed tomography images: A systematic review. Computer Methods and Programs in Biomedicine, 157, 49–67. https://doi.org/10.1016/j.cmpb.2018.01.014

[18] Baghdadi, N. A., Malki, A., Abdelaliem, S. F., Magdy Balaha, H., Badawy, M., & Elhosseini, M. (2022, May). An automated diagnosis and classification of COVID-19 from chest CT images using a transfer learning-based convolutional neural network. Computers in Biology and Medicine, 144, 105383. https://doi.org/10.1016/j.compbiomed.2022.105383

[19] Sudharson, S., & Kokil, P. (2020). An ensemble of deep neural networks for kidney ultrasound image classification. Computer Methods and Programs in Biomedicine, 197, 105709. https://doi.org/10.1016/j.cmpb.2020.105709

PLAGIARISM REPORT

201-15-14250

ORIGINALITY REPORT

23%	19%	14%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	dspace.daffodilvarsity.edu.bd:8080 Internet Source	4%
2	Submitted to Daffodil International University Student Paper	3%
3	Submitted to CSU Northridge Student Paper	1%
4	iieta.org Internet Source	1%
5	R. N. V. Jagan Mohan, B. H. V. S. Rama Krishnam Raju, V. Chandra Sekhar, T. V. K. P. Prasad. "Algorithms in Advanced Artificial Intelligence - Proceedings of International Conference on Algorithms in Advanced Artificial Intelligence (ICAAAI-2024)", CRC Press, 2025 Publication	1%
6	www.dlsu.edu.ph Internet Source	1%
7	www.frontiersin.org Internet Source	1%
8	spu-edu-krd.spu.edu.iq Internet Source	1%
9	Submitted to University of Essex Student Paper	<1%