



BSCS-S24-002

03-134211-022 Muhammad Ateeb Ather

Kvasir – Automated Gastro-Intestinal Disease Classification.

In partial fulfilment of the requirements for the degree of
Bachelor of Science in Computer Science

Supervisor: Abdullah

Department of Computer
Sciences Bahria University,
Lahore Campus

January 2025

C e r t i f i c a t e



We accept the work contained in the report titled
“Kvasir – Automated Gastro-Intestinal Disease Classification.”

written by

Muhammad Ateeb Ather

as confirmation to the required standard for the partial fulfilment of the degree of
Bachelor of Science in Computer Science.

Approved by:

Supervisor:

Abdullah

(Signature)

January 05, 2025

DECLARATION

We hereby declare that this project report is based on our original work except for citations and quotations which have been duly acknowledged. We also declare that it has not been previously and concurrently submitted for any other degree or award at Bahria University or other institutions.

| Enrolment | Name | Signature |
|---------------|----------------------|-----------|
| 03-134211-022 | Muhammad Ateeb Ather | |

Date: January 05, 2025

Especially dedicated to
My beloved father, mother and grandparents
(Muhammad Ateeb Ather)

ACKNOWLEDGEMENTS

I would like to extend my heartfelt thanks to everyone who played a part in the successful completion of this project. I am deeply grateful to my supervisor, Abdullah, for his invaluable guidance, insightful advice, and immense patience throughout the research process.

Additionally, I would like to express my sincere appreciation to my supportive parents and friends for their constant motivation and encouragement.

Muhammad Ateeb Ather

Kvasir – Automated Gastro-Intestinal Disease Classification.

Abstract

GI diseases are a major health concern in the world today and calling for prompt and accurate diagnosis for treatment. This study intends to capture the increasing incidence of GI diseases by devising an early detection and management mobile application using deep learning. It combines medical image analysis with patients' data to forecast GI conditions with relative accuracy, via features of image data. After testing the model for testing and evaluating different models of Deep learning, the best one is chosen for the diagnosis of GI diseases. The mobile application is intended to be very easy to use with an account creation page, a login page and a diagnosis page where uploaded medical images are sorted, classified, and predictions about potential GI diseases are made. Flask is used to develop the API to interconnect the mobile app and the deep learning model so that the real time predictions can be made as well as recommendations. The integration will help healthcare workers and clients get early warning signs to seek doctor's checkup in case of possible GI complications. By linking between the theoretical models and the application, this project aims at making a positive impact on the direction and efficient care of managing GI diseases and the patient's experience through it.

Table of Content

| | |
|---|-----------|
| Chapter 1 | 12 |
| Introduction | 12 |
| 1.1. Background | 12 |
| 1.2. Problem Statements | 13 |
| 1.3. Aims and objectives | 13 |
| 1.4. Scope of Project | 13 |
| 1.5. Kvasir Dataset | 14 |
| Chapter 2 | 15 |
| Literature Review | 15 |
| 2.1 Papers | 15 |
| Chapter 3 | 24 |
| Design and Methodology | 24 |
| 3.1. Dataset Collection and Preprocessing | 24 |
| 3.2. Data Augmentation | 24 |
| 3.3. Data Splitting | 24 |
| 3.4. Model Selection and Training | 25 |
| 3.5. Application of Boosting Algorithms | 25 |
| 3.6. Model Saving and Inference | 25 |
| 3.7. Design Diagrams | 26 |
| Chapter 4 | 34 |
| Implementation | 34 |
| 4.1. Mobile App Development | 34 |
| 4.2. Core Screens in the Mobile App | 34 |
| 4.3. User Registration and Validation | 35 |
| 4.4. API Integration with Flask | 35 |
| 4.5. Model Loading and Inference | 35 |
| 4.6. Tools and Technologies | 35 |
| 4.7. Implementation Process | 36 |
| Chapter 5 | 38 |
| Results And Discussions | 38 |
| 5.1. Performance Evaluation | 38 |
| 5.2. Confusion Matrix | 38 |
| 5.3. Splash Screen: | 39 |

| | |
|-------------------------------|-----------|
| 5.4. Registration Screen | 40 |
| 5.5. Login Screen: | 41 |
| 5.6. Diagnosis Screen: | 42 |
| 5.7. Result Screen: | 43 |
| Chapter 6 | 44 |
| Conclusion and Recommendation | 44 |
| 6.1. Conclusion | 44 |
| 6.2. Recommendation | 44 |
| References | 45 |

List of Tables

| | |
|---|----|
| Table 2.1.1: GI Abnormality Detection Using EWT and CNN | 17 |
| Table 2.1.2: DeepPoly for Polyp Segmentation and Classification | 19 |
| Table 2.1.3: Esophagitis Detection using Deep Learning Models | 19 |
| Table 2.1.4: Stomach Abnormality Detection Using Deep Learning | 20 |
| Table 2.1.5: GI Image classification Using Deep Learning architecture via Transfer Learning | 21 |
| Table 2.1.6: Spatial-Attention ConvMixer Architecture for GI Disease Detection | 22 |
| Table 2.1.7: A Robust Deep Learning Framework for Automated diagnosis of GI Diseases | 23 |
| Table 2.1.8: Performance Metrics across Dataset Categories and Models | 23 |

List of Figures

| | |
|--|----|
| Figure 3.7.1.1: Model Architecture Diagram | 26 |
| Figure 3.7.2.1: Application Flow Diagram | 27 |
| Figure 3.7.3.1 Activity Diagram | 28 |
| Figure 3.7.4.1: Data Flow Diagram Level 0 | 29 |
| Figure 3.7.5.1: Data Flow Diagram Level 1 | 30 |
| Figure 3.7.6.1 Data Flow Diagram Level 2 | 31 |
| Figure 3.7.7.1 Research Flow Diagram | 32 |
| Figure 3.7.8.1 Sequence Diagram | 33 |
| Figure 4.7.1.1: Phase 1: Core Screens and User Registration | 36 |
| Figure 4.7.2.1: Phase 2: API Integration, Model Loading, and Tools | 37 |
| Figure 5.1.1: Performance Evaluation | 38 |
| Figure 5.2.1: Confusion Matrix | 38 |
| Figure 5.3.1: Splash Screen: | 39 |
| Figure 5.4.1: Registration Screen | 40 |
| Figure 5.5.1: Login Screen | 41 |
| Figure 5.6.1: Diagnosis Screen | 42 |
| Figure 5.7.1: Result Screen | 43 |

Chapter 1

Introduction

1.1. Background

GI diseases are a major health problem and have a large impact on morbidity and healthcare cost. These diseases if diagnosed early and well managed have minimal risks of complications hence the need to have them diagnosed and treated. Many classification approaches that are presently used involve visual examinations by clinicians which are cumbersome and prone to errors. Therefore, the need for accurate diagnostic biomarkers for early detection of Gastrointestinal (GI) conditions using features such as image processing and deep learning algorithms is more and more felt. Computer vision helps to diagnose patients based on images as applied in dermatology and gastroenterology since patterns are critical in diagnose various diseases. However, with regards to GI diseases, the use of features has been credited with significant success in classifying endoscopic images). Using large data sets such as the Kvasir v2 dataset that has a large mix of GI disease images these models are well equipped to learn additional information that the clinicians on their own might not recognize or capture quickly or in the right manner.

This project based on the detection of GI disease history, which employs image processing and deep learning solutions for condition detection from the endoscopic images. Kvasir v2 contains more diverse images belonging to 8 classes of GI diseases. Different levels of preprocessing operations along with data augmentation and data balancing were done to improve the model performance. Several deep learning models including HLNN, VIT, Resnet50, Inception Resnet models were trained and tested, resulting in a tool that health care givers can use to diagnose the GI diseases more accurately and with more efficiency.

1.2. Problem Statements

GI diseases however are a major concern in the lives of affected individuals because they often lead to severe complications if not diagnosed in early stages and are expensive to manage. Current approaches in diagnosing GI conditions does not offer accurate evaluation and involve merely observation and physical assessment by clinicians. Since gastrointestinal diseases are on the rise, the current gap has a significant need for an automated, accurate, and efficient algorithm to assist early detection and enhance patient care and healthcare systems.

1.3. Aims and objectives

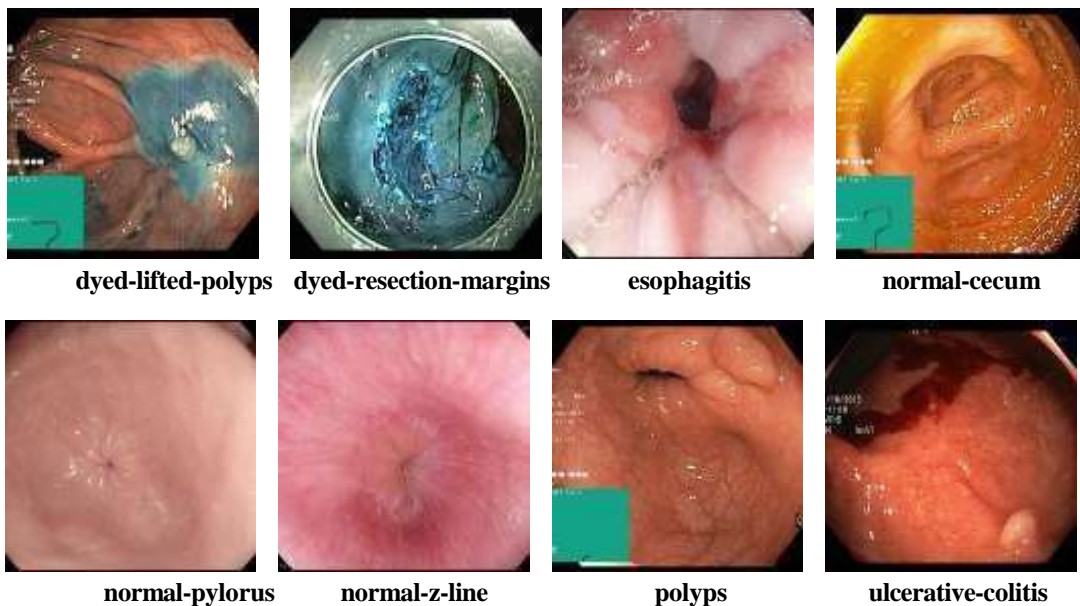
The objective of this project is thus to design a web-based Application, specifically a GI disease Detection Tool that uses some deep learning algorithm to diagnose the disease when an image that represents an affected GI organ is uploaded. The objectives include: (1) collecting and preprocessing Kvasir dataset of GI disease images, (2) applying image augmentation and balancing techniques to enhance the dataset, (3) training and evaluating various deep learning models such as Hlnn, Vit, Resnet and EfficientNetB0, and (4) deploying the final model into a mobile application for easy accessibility by healthcare professionals.

1.4. Scope of Project

Specifically, this project focuses at designing an image based diagnostic system for GI diseases using machine learning algorithms. It is the process of enhancing the probabilistic outcomes of deep learning application, image classification. The tool is intended for use in clinical settings and aims to assist healthcare professionals in making timely and accurate diagnoses. The project also includes the deployment of the model within a mobile application, making it accessible for use in real-time patient care scenarios.

1.5. Kvasir Dataset

Kvasir is the large dataset of annotated medical images, which are approved by experienced endoscopists and can be used for image search, machine and deep learning as well as transfer learning. This includes quality images (720×576 to 1920×1072) across the different classes of anatomical landmarks, pathological findings and endoscopic procedures within the gastrointestinal (GI) tract. Findings mentioned as Z-line, pylorus, cecum assist with navigation while inspecting the inside of the digestive tract as well as help determine diseases like Gastro- esophageal reflux or ulcers. Esophagitis, polyps, ulcers are examples of features which can be visualized and correlated to the severity and progression of the disease. The dataset also consists of pictures of endoscopic mucosal resection polypectomy and corresponding landmarks including dyed and lifted polyps and resection margins to allow for proper identification and evaluation. Some of the images also contain an insert illustrating the position of the endoscope, which provides valuable interpretative information.



This dataset is useful in identifying a Computer-aided detection schemes for diseases diagnosis, classification, and reporting that enables quick diagnosis, time saving and hence reduces the man hours, cost effectiveness and fully tasting the benefit of the patient. Hence, it can be used for benchmarking and replication and other studied object detection experiments, video analysis, and multimedia researches for the progress of health care technology and GI disease management.

Chapter 2

Literature Review

2.1 Papers:

The identification and categorization of gastrointestinal (GI) pathologies have come a long way due to newer developments in deep learning and signal processing. These technologies are driving the future of medical imaging because they facilitate effective diagnostic mechanisms. This has been done by combining Empirical Wavelet Transform (EWT) for feature extraction, and deep convolutional neural networks (CNNs) for classification. As a result of this, EWT is able to separate fine, frequency-like characteristics from images hence offering a strong feature map. CNNs then optimize these features to perform precise identification of the abnormalities, a step closer towards achieving compliant diagnostic aid for clinicians [1]. In the domain of CRC screening, DeepPoly can present a state-of-art solution with a kind of automation of the polyp segmentation and classification. By using the DoubleU-Net model, a high segmentation performance of a dice coefficient of 0.834 for the Endotech dataset and a dice coefficient of 0.956 for the Kvasir-SEG dataset is attained. The DeepPoly algorithm is complemented by the Vision Transformer (ViT) for classification of polyps into hyperplastic or adenomatous types with near perfect accuracy of 99%. This advancement enables identification of patients at high risk and could prevent CRC and promote better patient care. When combining accurate segmentation and resilience classification, deep learning intact is shown to reshape the approach to autonomous colonoscopy examinations [2]. At the same time, attempts to identify esophagitis – a disease that, if unnoticed, can become cancerous and affect the esophagus – have concentrated on comparing various types of deep learning. Authors analyzed existing structures with defined quality on the Kvasir dataset for architectures such as GoogLeNet, ResNet-50, MobileNet V2, and MobileNet V3. Of the four models, GoogLeNet had the best series of F1-scores, whereas MobileNet V3 was more confident in the predictions. For improvement of interpretability the methods which were used are SHapley Additive exPlanations (SHAP) and

Gradient-weighted Class Activation Mapping (Grad-CAM), which give a deeper understanding of how such models come to their findings. Such transparency is particularly beneficial when clinical software is developed and used in applications that require diagnostics based on the decision-making of the artificial intelligence, [3] where patients rely on the credibility of a diagnostic algorithm. There is another innovative method which distinguishes and characterizes GI abnormalities following a step-by-step procedure. First, based on threshold of RGB channel after pixel- based techniques and mutual information, the location of interest regions, for instance, bleeding areas are segmented. These segmented regions are then characterized and discriminated using texture obtained from Local Binary Pattern (LBP) and deep features acquired using transfer learning. Features are then selected using entropy and the results are demonstrated through an Ensemble Learning Classifier. The applied method was 99.8% accurate with a private dataset and 86.4% with Kvasir, indicating its applicability to various clinical cases [4]. Specifically, transfer learning, as one of the most sought-after techniques in the currently developing AI, has been crucial for the further enhancement of GI image classification. Using transfer learning where massive pre-trained models are fine-tuned solves the above problem, which is a typical inconvenience in medical imaging—limited labeled datasets. Comparing several architectures namely VGG19, ResNet50V2, ResNet152V2, EfficientNetV2B0, EfficientNetV2B3, InceptionV3, DenseNet201, Xception outperformed all the other models where the VGG19 achieved a maximum of 88.6% classification accuracy. This goes to show that transfer learning is effective in enhancing the classification model that is to be used in advancing medical field as it is precision formulated to be clinically effective [5]. New architectures are also being developed to overcome the primary challenges of GI disease detection. Notable among these is the Spatial-Attention ConvMixer (SAC) model which includes novel spatial attention mechanisms into the ConvMixer model. This addition empowers the model to attend to essential sections in an image; an ability helpful in the identification of abnormalities. In terms of accuracy, SAC obtained an outstanding yield of 93.37% on the Kvasir dataset outcomping with other more conventional models such as ViT, Swin Transformer, ResNet50, and more. The study shows that SAC can enhance GI disease diagnosis and provides a promising approach to currently existing techniques [6]. Besides these architectural changes, the improvements in data enhancement have also amplified the

diagnostic models. For example, the researchers using EfficientNetB5 architecture used horizontal flipping and dynamic scaling to reduce variability in medical image data sets. This approach gave a test accuracy of 98.89% and a test loss of 0.230 thereby showing that the model was capable of good generalization without having overly high levels of overfitting. These discoveries highlight the need for incorporating high-performing architectures with legitimate augmentation approaches as a means of strengthening the diagnostic systems [7]. Taken together, all these advances bear witness on the ability of deep learning in revolutionizing medical imaging domain. Through methods like EWT, spatial attention mechanisms, transfer learning, and interpretability tools, the machine learning and Deep Learning researchers have made models that can give high accuracy, and at the same time, report meaningful implications for the healthcare practitioners. When segmentation and classification are integrated into a pipeline, the capability of such systems is further improved in terms of both coverage and computational costs for complicated diagnostic tasks. Which brings one back to the understanding of why explain ability is such a key in AI-systems when you are within the medical field. Some of AI explanations include SHAP and Grad-CAM which help the healthcare professionals to be able to read onto the reasoning of the model and hence being able to trust on AI diagnosis. In the clinical practice context, where consequence implications are strong, such interpretability is desired to support decision-making dependability and accuracy. The presence of both segmentation and classification pipelines that are evident in DeepPoly and other methods are an integration of a comprehensive solution towards screening of GI disease. Segmentation we separate the regions of interest that may include the polyps or the bleeding regions while classification tells the kind of defect. This two-path approach does not only enhance the accuracy of the diagnoses but also eases clinicians' decision tasks, thus rendering AI implementation tools more feasible and effective in actual practice. Using transfer learning remains the cornerstone of advancing this line of development. It has been observed that by fine tuning some existing models, with new datasets, much has been accomplished even with the data limitation. It is moreover shown by architectures such as VGG19 and EfficientNetB5 that by excellent implementation of transfer learning along with adequate augmentation, AI models are capable of achieving highest level of accuracy and transferability. New architectures such as SAC represents the future of

medical AI. These models tailored for GI disease detection introduce domain-specific improvements which includes spatial attention to handle detailed difficulties. Kvasir dataset is the benchmark case where such innovations like SAC have the potential to achieve improved accuracy and dependable diagnostic performance. Thus, the combination of latest deep learning methods with medical imaging offers a new promising approach to detect and classify the GI abnormalities. From using EWT for extracting features to integrating transfer learning and state of the art architectures these are the means that are exerting advancement pressure in the realms of automated diagnostics. The incorporation of these DCIPs including segmentation-classification pipelines, interpretability tools, as well as advanced data augmentation guarantees that such models are accurate as well as feasible and reliable. With the further development of these technologies, good potential for improving diagnostic activity, minimizing the load on clinicians, and, as a consequence, the incentives to achieve better results in treating patients can be outlined.

Table 2.1.1: GI Abnormality Detection Using EWT and CNN

| | |
|---------------------------|--|
| Article Name: | Gastrointestinal abnormality detection and classification using empirical wavelet transform and deep convolutional neural network from endoscopic images |
| Problem Statement: | The differentiation of GI abnormalities in endoscopic images is challenging because of the stability of the image quality and diseases featured by their appearance and structural characteristics. Here, that means that the variations present difficulties for conventional wisdom over how precisely they should be managed. |
| Model/Technique: | Dataset Augmentation Empirical Wavelet Transform (EWT) Inception CNN |
| Results: | For the classification of GI abnormalities the proposed model achieved 96.5% accuracy. As for endoscopic image analysis, it was demonstrated that it surpassed traditional image analysis methods in both accuracy and resistance to distinct types of abnormalities. |

| | |
|-----------------|---|
| Gaps: | Need for enhancement in handling diverse datasets and real-world clinical conditions. |
| Classes: | The model classifies multiple types of gastrointestinal abnormalities, which include but are not limited to polyps, ulcers, inflammatory diseases, and other lesions detected in endoscopic examinations. |

Table 2.1.2: DeepPoly for Polyp Segmentation and Classification

| | |
|---------------------------|--|
| Article Name: | DeepPoly: Deep Learning-Based Polyps Segmentation and Classification for Autonomous Colonoscopy Examination |
| Problem Statement: | The challenge of accurate segmentation and classification of polyps in colonoscopy images to assist in the early detection of colorectal cancer. |
| Model/Technique: | DeepPoly applies deep learning methodologies The proposed method utilized DoubleU-Net for polyps segmentation and Vision Transformer (ViT) for risk classification of the polyps. |
| Results: | The method accurately reached a 99% performance and accuracy rate for polyp’s detection and classification and has the potential for further use in actual colonoscopy examinations. |
| Gaps: | The model requires further optimization for handling diverse datasets and improving real-time processing capabilities in clinical settings. |
| Classes: | The model classifies different types of polyps (e.g., adenomas, and hyperplastic polyps) as part of its segmentation and classification task. |

Table 2.1.3: Esophagitis Detection using Deep Learning Models

| | |
|---------------------------|--|
| Article Name: | Deep-learning models in medical image analysis: Detection of esophagitis from the Kvasir Dataset |
| Problem Statement: | Automating diagnosis of esophagitis from endoscopic images is also not easy since there is variation in quality and appearance of images. |
| Model/Technique: | In the systematic review, the article uses deep learning models, namely CNNs CNVs, to assess images from the Kvasir dataset for esophagitis Pre-processing and transfer learning methods are applied to improve model efficiency. |
| Results: | The deep learning model demonstrated strong performance, achieving significant classification accuracy for detecting esophagitis in the Kvasir dataset, surpassing traditional methods. GoogLeNet : 84.6%, ResNet-50 : 83.3%, MobileNet V2 : 83.0%, and MobileNet V3 84.2% |
| Gaps: | Further improvements are needed for generalization to diverse datasets, handling different clinical environments, and improving the model's robustness to noisy or low-quality images. |
| Classes: | The model classifies images into two categories, focusing on the detection of esophagitis and z-line in endoscopic images with potential extensions to other GI abnormalities. |

Table 2.1.4: Stomach Abnormality Detection Using Deep Learning

| | |
|---------------------------|---|
| Article Name: | Segmentation and Classification of Stomach Abnormalities Using Deep Learning |
| Problem Statement: | Detecting and classifying stomach abnormalities from medical images is challenging due to varying image quality, abnormal appearance, and overlapping features with other conditions. |

| | |
|-------------------------|--|
| Model/Technique: | The system uses a two-step pipeline: segmentation of the bleeding infection region using a hybrid method and classification using deep learning with transfer learning (TL) for feature extraction. Local Binary Pattern (LBP) is used for texture features, and an entropy-based feature selection approach is applied before feeding the optimal features to an ensemble classifier. |
| Results: | The system achieved 99.8% accuracy on a private dataset and 86.4% accuracy on the KVASIR dataset, outperforming other methods in classifying GI abnormalities. |
| Gaps: | Further work is required to address the model's adaptability to diverse datasets and the generalization to real-world, clinically diverse scenarios. |
| Classes: | The model detects various gastrointestinal abnormalities, including bleeding, polyps, and ulcers. |

Table 2.1.5: GI Image Classification via Transfer Learning

| | |
|---------------------------|--|
| Article Name: | Gastrointestinal Image Classification Using Deep Learning Architectures via Transfer Learning |
| Problem Statement: | Detecting and classifying gastrointestinal diseases using medical images is challenging, requiring accurate and efficient systems to aid medical professionals. |
| Model/Technique: | The study uses transfer learning with Convolutional Neural Network (CNN) architectures including VGG19, ResNet50V2, ResNet152V2, EfficientNetV2B0, EfficientNetV2B3, InceptionV3, DenseNet201, and Xception for classification of GI images from the Kvasir dataset. |
| Results: | Among all established models, VGG19 emerged as the best with an accuracy of 88.6% illustrating that it was better placed at distinguishing gastrointestinal images. |

| | |
|----------------|--|
| Gaps: | Potential areas for improvement include refining model performance across all architectures and exploring more diverse datasets for better generalization. |
| Classes | The models are able to categorize various gastrointestinal pathologies including dyed-lifted polyps, dyed-resection margins, esophagitis, normal caecum, normal pylorus, normal z-line, polyps and ulcerative colitis. |

Table 2.1.6: Spatial-Attention ConvMixer architecture for GI Disease Detection

| | |
|---------------------------|--|
| Article Name: | Spatial-attention ConvMixer architecture for classification and detection of gastrointestinal diseases |
| Problem Statement: | Manual evaluation of GI disorders can be error-prone, and the high similarity between classes makes automated classification challenging. |
| Model/Technique: | SAC model operates at integrating ConvMixer with the proposed Spatial Attention Mechanism (SAM). |
| Results: | Unlike Vanilla ViT, Swin Transformer, ConvMixer, MLP Mixer, ResNet50, and SqueezeNet, proposed SAC method attains 93.37% accuracy, while the other architectures have respectively achieved only 79.52%, 74.52%, 92.48%, 63.04%, 87.44% and 85.59% accuracy. |
| Gaps: | Despite SAC's strong performance, there's room for further improvement in generalizing across diverse datasets. |
| Classes: | The Kvasir dataset includes 8 classes |

Table 2.1.7: A Robust Deep Learning Framework for Automated Diagnosis of GI Diseases

| | |
|---------------------------|---|
| Article Name: | Enhancing image-based diagnosis of gastrointestinal tract diseases through deep learning with EfficientNet and advanced data augmentation techniques |
| Problem Statement: | Traditional models for medical imaging diagnosis suffer from overfitting and lack generalization due to the variability and complexity of disease manifestations. |
| Model/Technique: | EfficientNetB5 architecture combined with transfer learning, maximal pooling, advanced regularization, and sophisticated data augmentation techniques (e.g., horizontal flipping, dynamic scaling). |
| Results: | Attained 98.89 % test accuracy, 0.230 test loss and all classes' high precision as compared to classical techniques. |
| Gaps: | These models only include Limited data augmentation and regularization mechanisms so they fail to give proper results in composite real-world changing data sets. |
| Classes: | Ulcerative colitis, polyps, esophagitis, and other gastrointestinal diseases. |

Table 2.1.8: Performance Metrics across Dataset Categories and Models

| Datasets | Model Techniques | Methods | Accuracy | F1 | Recall | Precision |
|-----------------------|-------------------------------|------------------|-----------------|-----------|---------------|------------------|
| Kvasir Dataset | HLNN | train test split | 81.5 | 81.02 | 80.9 | 81 |
| Kvasir Dataset | VIT | train test split | 92 | 91.36 | 90.38 | 91.50 |
| Kvasir Dataset | Hlnn+Vit with meta-classifier | train test split | 91 | 91 | 91 | 91 |

Chapter 3

Design and Methodology

3.1. Dataset Collection and Preprocessing

- The dataset consists of gastrointestinal (GI) disease images categorized into eight classes: 'dyed-lifted-polyps', 'dyed-resection-margins', 'esophagitis', 'normal-cecum' , 'normal- pylorus' , 'normal-z-line' , 'polyps' , 'ulcerative-colitis'.
- Data preprocessing includes resizing, rescaling images, normalization, and other techniques to improve data quality and consistency. Also keeping in mind to maintain the aspect ratio while resizing to avoid distortion and loss of features.

3.2. Data Augmentation

- The Data augmentation techniques are employed to increase the size of the dataset and on the basis of the quality of augmented images, Mean Squared Error (MSE), Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR) are used. This means the augmented images are not significantly different from the original images and also enhances the robustness of the model since the data integrity is preserved while increasing the data size in the training set.”

3.3. Data Splitting

- The dataset is split into three parts: 70% for training, 15% for testing, and 15% for validation.

3.4. Model Selection and Training

- HLNN (Hybrid Learning Neural Network), a cutting-edge hybrid model combining multiple neural network architectures, is selected for model training. The model is trained for 50 epochs, with performance monitored during the process.
-
- ViT (Vision Transformer), a state-of-the-art transformer-based architecture, is selected for model training.

3.5. Application of Boosting Algorithms

- After achieving initial accuracy both the model, boosting algorithms XGBoost is applied to enhance the hybrid model's predictive accuracy.

3.6. Model Saving and Inference

- The trained model, achieving 94% accuracy during training and 91% accuracy during testing, is saved in an .h5 file for deployment.
- A prediction function is developed to test the model and assess its performance.

3.7. Design Diagrams

3.7.1. Model Architecture Diagram

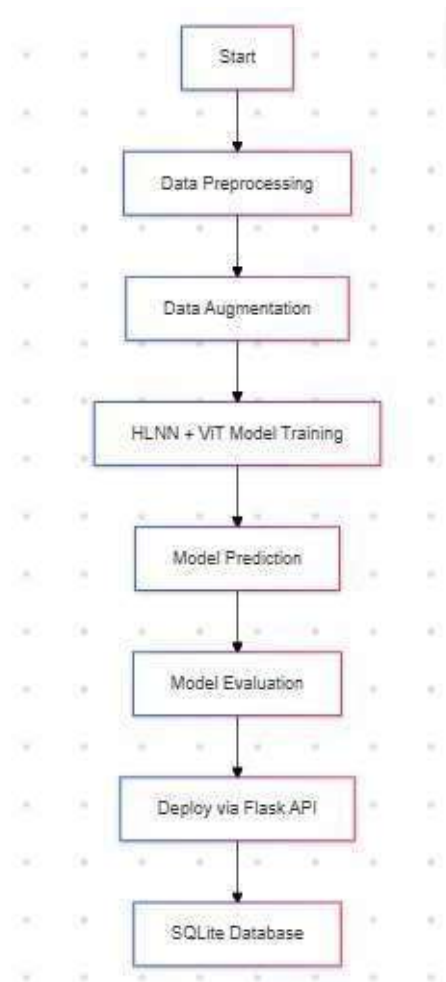


Figure 3.7.1.1: Model Architecture Diagram

The architectural model involves a doctor or diagnostician working alongside the stages of a Deep Learning model for GI disease classification. The model begins with data preprocessing, then uses a Vision Transformer (ViT) for feature extraction, and ends with a classification layer that outputs the predicted class labels. The possible output classes include 'dyed-lifted-polyps,' 'dyed-resection- margins,' 'esophagitis,' 'normal-cecum,' 'normal-pylorus,' 'normal-z-line,' 'polyps,' and 'ulcerative- colitis.' The final class label is used to offer relevant suggestions to users.

3.7.2. Application Flow Diagram

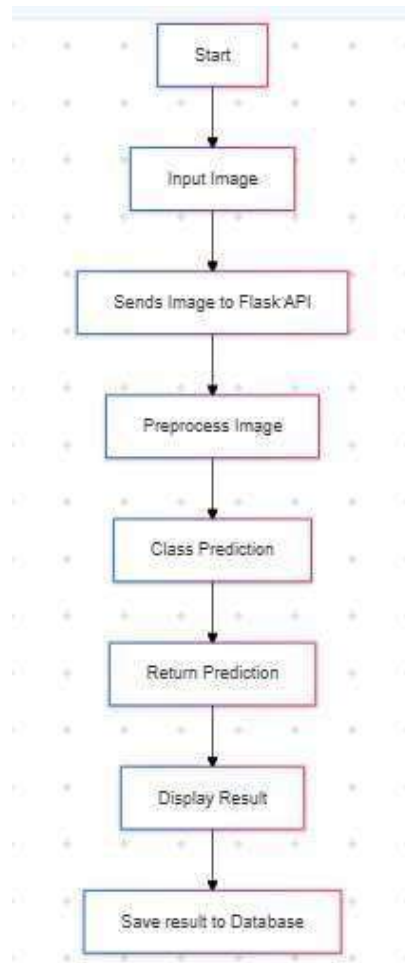


Figure 3.7.2.1: Application Flow Diagram

The application flow diagram shows the functionality of the mobile app by showing the steps to be followed in the application. It then extends to user's registration and authentication, then image upload for analysis. When the image is uploaded, the results are obtained via an API so as to get the results of the deep learning model. The last step involves the display of the result; the predicted class of the image and some recommendations in relation to the user.

3.7.3. Activity Diagram

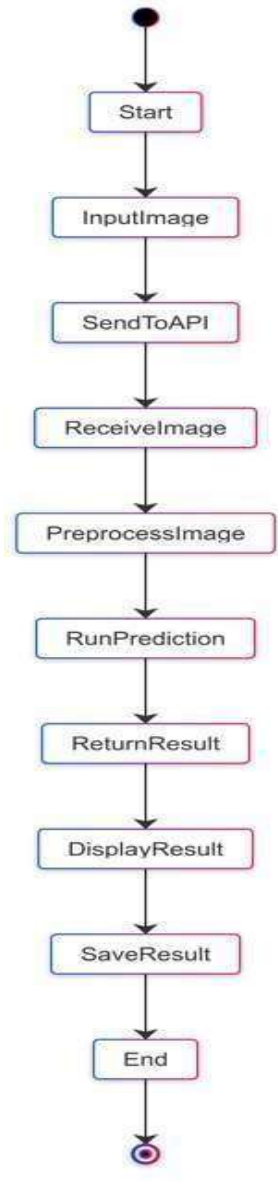


Figure 3.7.3.1 Activity Diagram

The activity diagram signifies the sequence of operations within the mobile app, operations such as registration and log in of the user. Having entered all the necessary data and passed the authentication, the user takes a picture of his or her foot and uploads it into the computer; a complex of algorithms with the help of artificial neural networks classifies the picture. According to the model the following is displayed as the predicted. If the user is not registered he simply completes his registration process the program guides him/she.

3.7.4. Data Flow Diagram Level 0

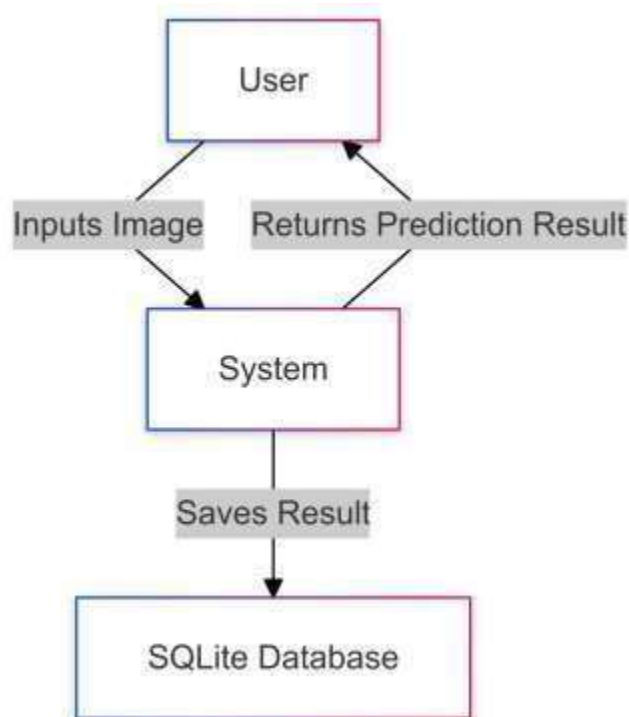


Figure 3.7.4.1: Data Flow Diagram Level 0

This level offers an overview of the working of the system and on this level the process flow is generalized. It shows the main entities: The user model, the mobile application, and the deep learning model as well as the relationships between them. The foot images to be examined are uploaded by the user through the app and the app further interacts with the model for classification. The model shares the forecast with the user, and the app shows the prediction to the user as well.

3.7.5. Data Flow Diagram Level 1

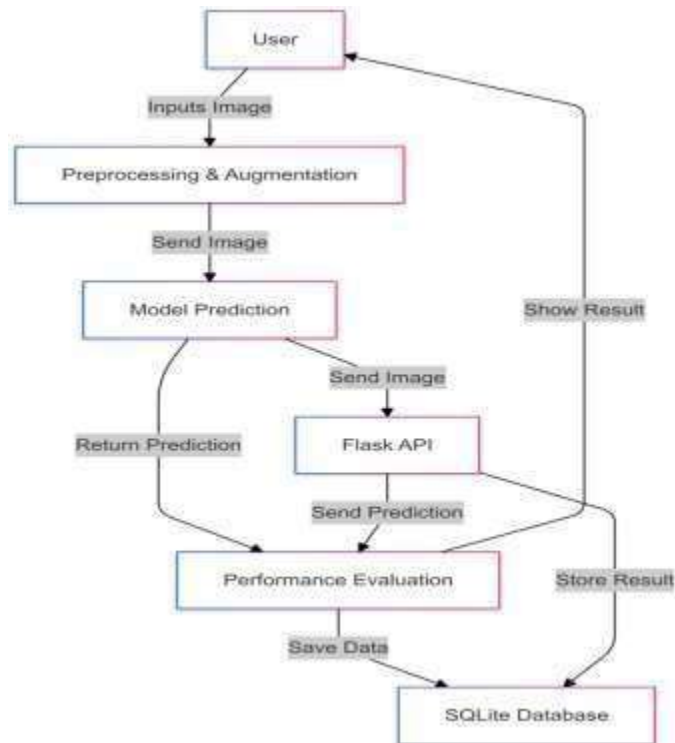


Figure 3.7.5.1: Data Flow Diagram Level 1

At this level, the diagram gets more specific on how the user is going to interact with the system. It includes user registration, login, image upload and classification. After the app captures an image it forwards image data to the deep learning model and that involves the process of analyzing an image to come up with a prediction of its class. The result is returned by this system to the app and is then shown to the user.

3.7.6. Data Flow Diagram Level 2

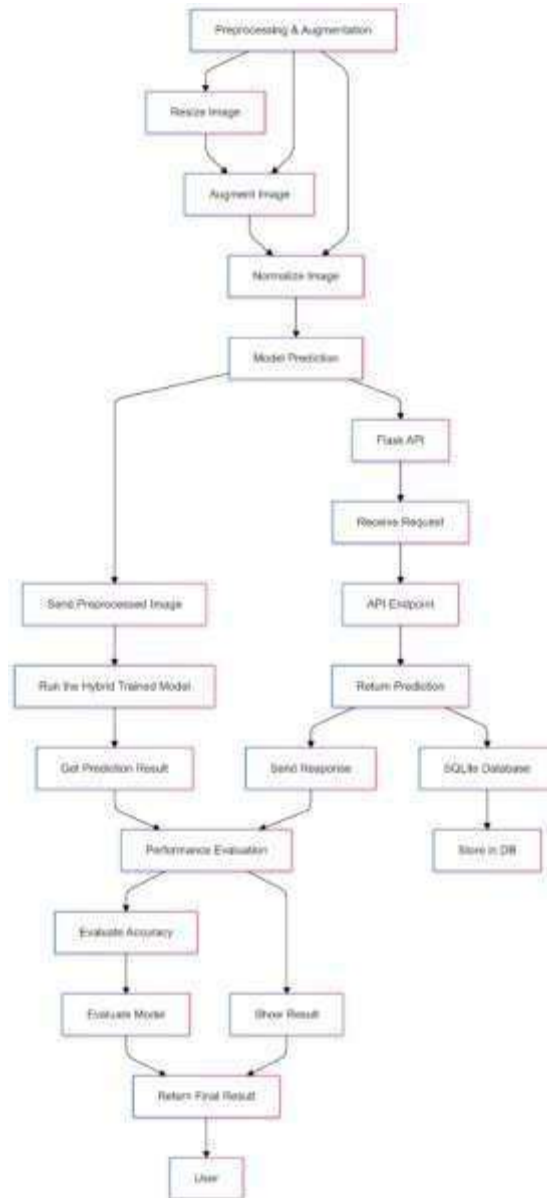


Figure 3.7.6.1 Data Flow Diagram Level 2

Level 2 introduces additional internal process information of higher granularity than the information at Level 1. Within it, process and activity of the mobile application are presented for example, account authorization, email and password validation, image size rendition and normalization, API cooperation with Flask. The diagram also address the workflow of the deep learning model inside the total computational system, which includes the model's component where the model performs the classification and returns the result back to the app processing it before presenting to the end user.

3.7.7. Research Flow Diagram

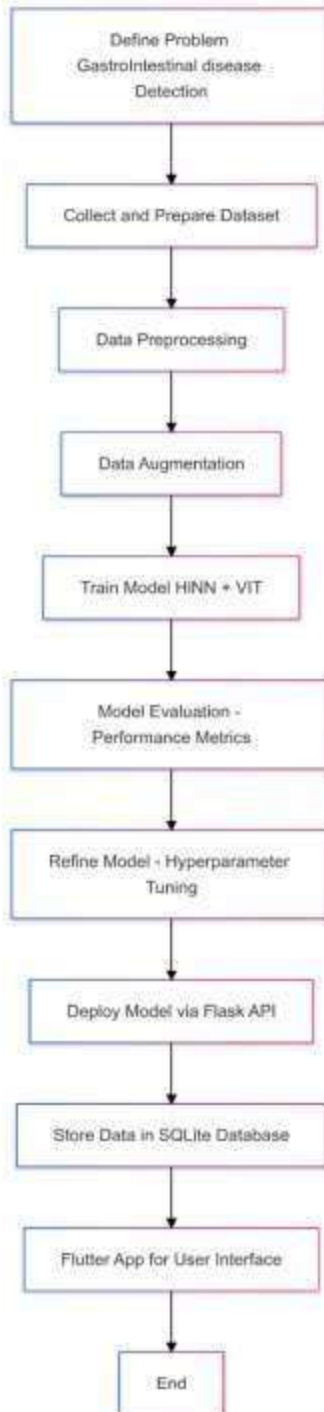


Figure 3.7.7.1 Research Flow Diagram

This flowchart shows the steps right from data collection to deployment of the final model. Next, features are extracted from the preprocessed dataset which is again a fine-

tuned dataset of diabetic foot ulcer images. When the data is divided into training, testing and validation data, the HInn Vit model is trained. After satisfactory results, the created model is stored and integrated into a function in a mobile application. This is because the app enables user to upload images in order for the model to predict the given class of the image and the recommendation.

3.7.8. Sequence Diagram:

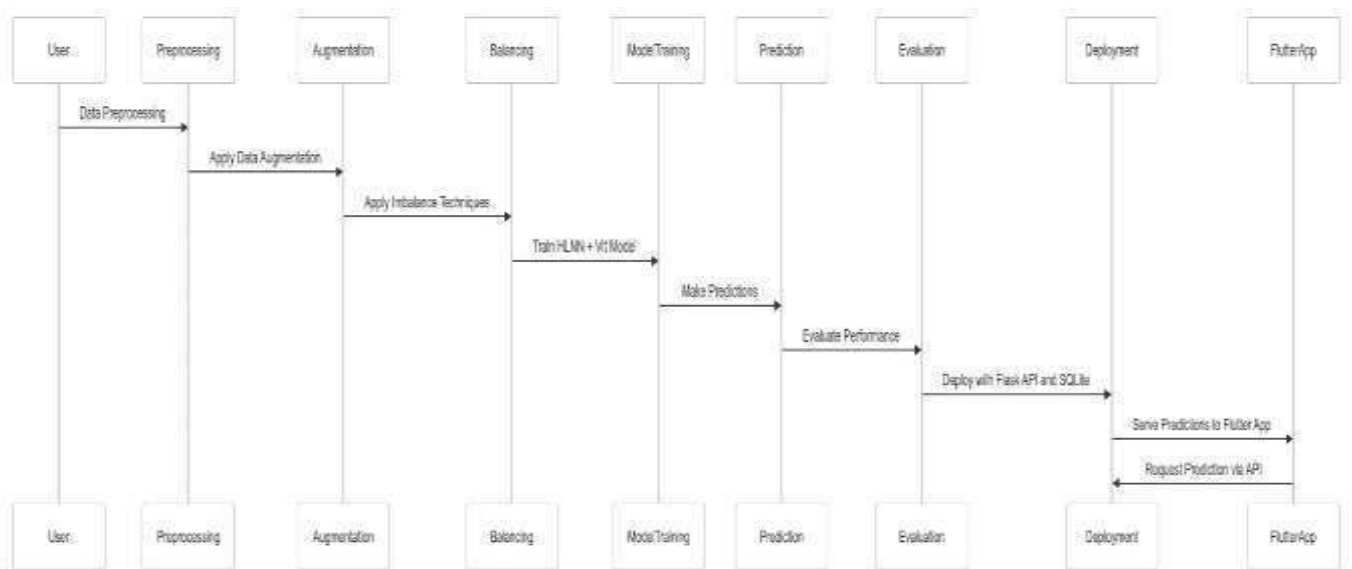


Figure 3.7.8.1 Sequence Diagram

The flow of conversation among user, mobile app and the deep learning model is displayed in the sequence diagram. It begins with a user signing up, or even if they already have an account, they sign into the app to post a GI image for analysis. The app is then used to call the model using an API to predict the class of the image that was uploaded. The result is consequently sent back from the model to the application and shown to the user.

Chapter 4

Implementation

4.1. Mobile App Development

- The mobile app is developed using Flutter for cross-platform compatibility and Android Studio as the IDE.

4.2. Core Screens in the Mobile App

4.2.1. Splash Screen

- Displays the app logo briefly.

4.2.2. Registration Screen

- Users enter their email, password, and confirm password with validation checks.

4.2.3. Login Screen

- Registered users can log in with their credentials.

4.2.4. Diagnoses Screen

- Users upload gastrointestinal images, which are analyzed by the model to predict the class from Kvasir dataset and provide recommendations.

4.3. User Registration and Validation

- Registration includes email validation and password length check, and the system checks
- Registration consist of email validation and password length check and the system check if the user already exists. First time users are required to create a new account.

4.4. API Integration with Flask

- A Flask-based API facilitates communication between the mobile app and the trained deep learning model, allowing for real-time predictions.

4.5. Model Loading and Inference

- The trained model, stored in an .h5 file, is loaded into the app to make predictions based on user-uploaded foot images.

4.6. Tools and Technologies

4.6.1. Programming Languages

- Python (for model development) and Flutter (for mobile app development).

4.6.2. Libraries

- TensorFlow, Scikit-learn, Pandas, NumPy etc.

4.6.3. IDE

- VS Code (for model development) and Android Studio (for mobile app development).

4.6.4. Database

- Used to store user information securely.

4.7. Implementation Process

4.7.1. Phase 1: Core Screens and User Registration

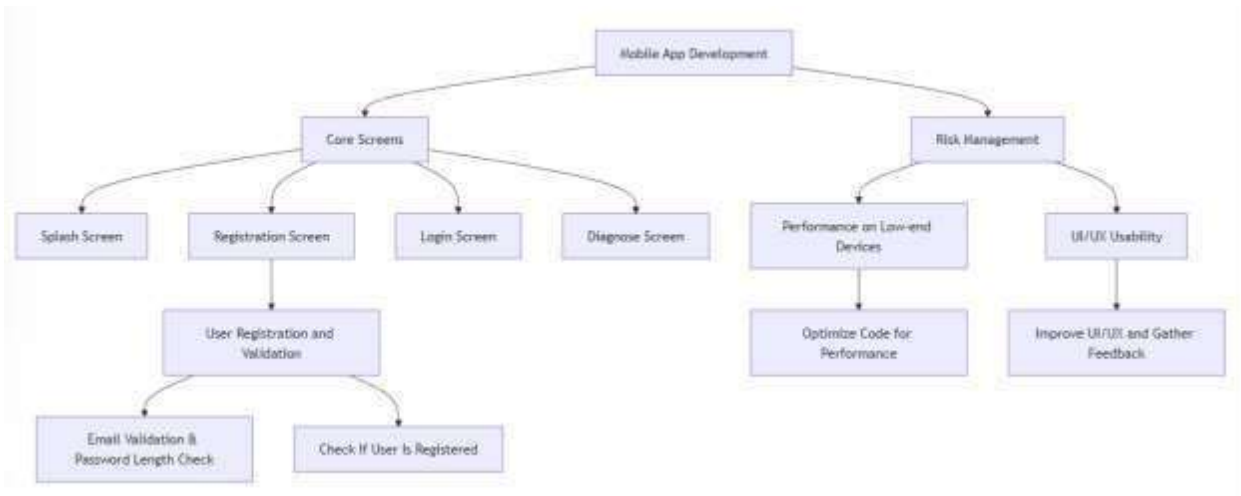


Figure 4.7.1.1: Phase 1: Core Screens and User Registration

This flowchart illustrates the main steps in model deployment and the working process of an app. The authors explain how exactly the user engages the given app, from creating an account to entering the credentials, and uploading images. After that, the app sends the image to a trained deep learning model via API to predict the class of the image, and then show the result.

4.7.2. Phase 2: API Integration, Model Loading, and Tools

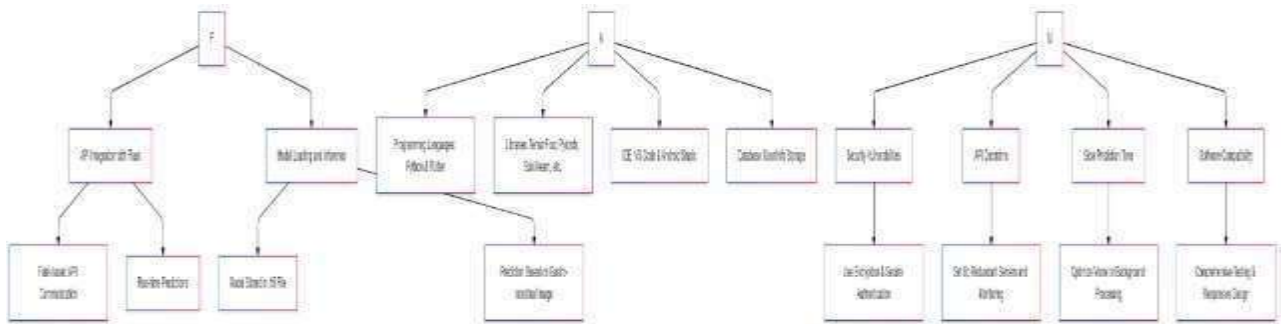


Figure 4.7.2.1: Phase 2: API Integration, Model Loading, and Tools

The following diagram shows the message flow in a broad level from the user inputs to the model and the output feedback. It demonstrates how the user posts an image and how through an API the image will go through the trained model. The model takes the image and processes it, and then the system gives out the classification it did on the image. The user interface is aimed at providing proper interaction and, in turn, the exchange of information between the app and the model.

Chapter 5

Results and Discussions

5.1. Performance Evaluation

| | precision | recall | f1-score | support |
|------------------------|-----------|--------|----------|---------|
| dyed-lifted-polyps | 0.91 | 0.89 | 0.90 | 150 |
| dyed-resection-margins | 0.91 | 0.93 | 0.92 | 150 |
| esophagitis | 0.85 | 0.77 | 0.80 | 150 |
| normal-cecum | 0.93 | 0.95 | 0.94 | 150 |
| normal-pylorus | 0.99 | 1.00 | 0.99 | 150 |
| normal-z-line | 0.79 | 0.85 | 0.82 | 150 |
| polyps | 0.93 | 0.90 | 0.92 | 150 |
| ulcerative-colitis | 0.97 | 0.97 | 0.97 | 150 |
| accuracy | | | 0.91 | 1200 |
| macro avg | 0.91 | 0.91 | 0.91 | 1200 |
| weighted avg | 0.91 | 0.91 | 0.91 | 1200 |

Figure 5.1.1: Performance Evaluation

5.2. Confusion Matrix

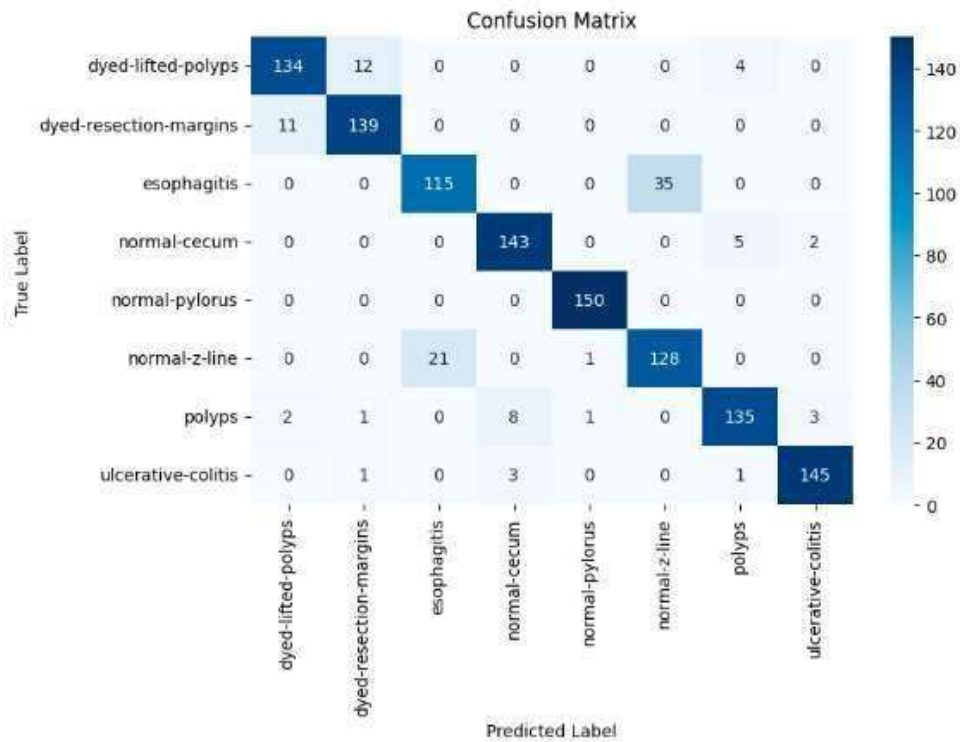


Figure 4.7.2.1: Confusion Matrix

5.3. Splash Screen:



Figure 5.3.1: Splash Screen:

The splash screen shows the symbol of the particular application just for a few seconds and gives the first positive impression to the user. It can also act as a front-end interface before getting the deep into the operational aspect of the application.

5.4. Registration Screen

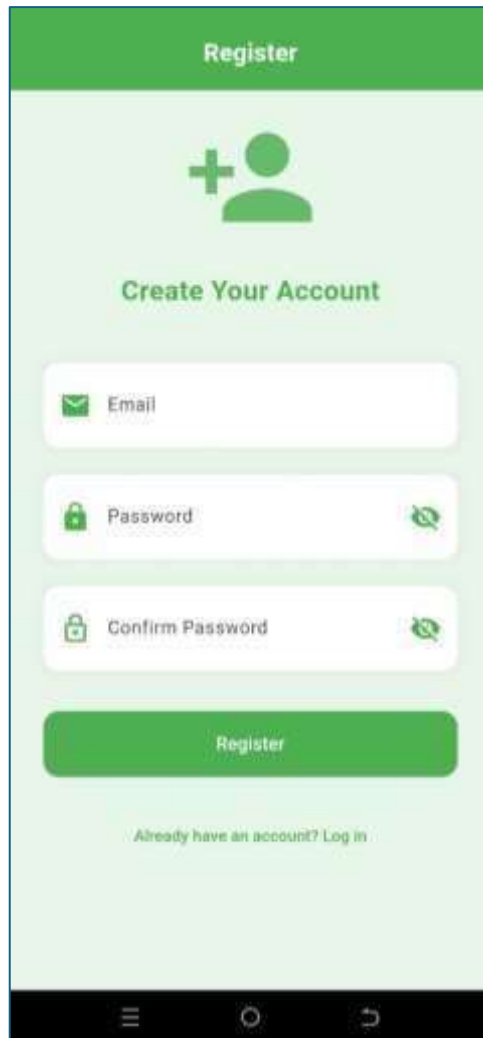


Figure 4.7.2.1 Registration Screen

The member's area is accessed by providing the login credentials – the user's email and password. If the login details are correct, then the users are enabled to the usage of the major functions of the app, otherwise they are asked to try again.

5.5. Login Screen:

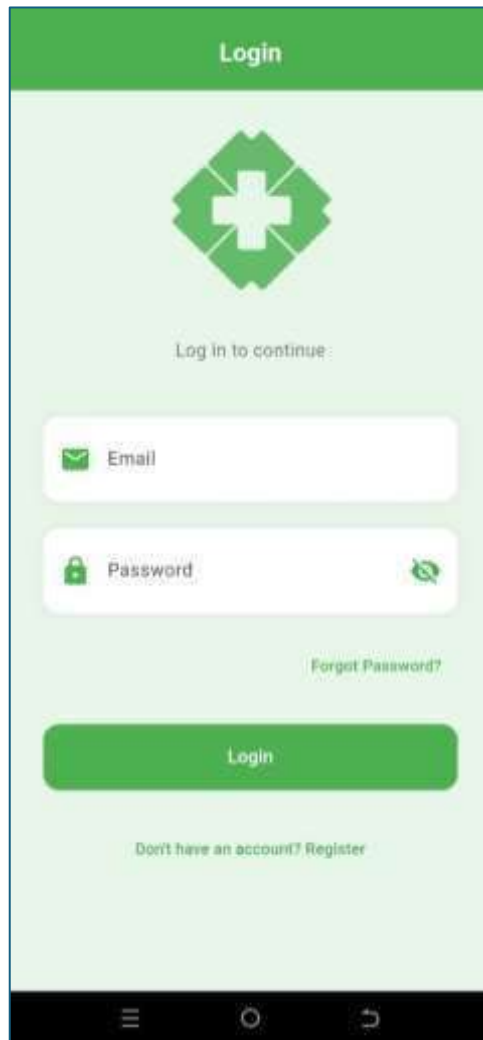


Figure 5.5.1 : Login Screen

It contains email, password and confirm password input fields, which first time users or any people who still do not have an account have to provide while registering On the registration screen. The app also implements validation checks preferring whether the inputs meet the following conditions, the correct format of an email and the length of the password. When a user comes to the page, if he/ she is registered then the user will be automatically redirected to the login page.

5.6. Diagnosis Screen:



Figure 4.7.2.1: Diagnosis Screen

The diagnose screen allows users to select a GI image to be analyzed. The picture is analyzed by the model to rank it into one of the 8 GI disease types. Finally, all the results concerning the user are shown.

5.7. Result Screen:



Figure 4.7.2.1: Result Screen

The result screen of diagnostic API shows the predicted class of the uploaded image below the analyzed image and suggests further steps or treatment. It also offers information about the extent of the disease and its type to the users and make recommendations on the management of the condition or its monitoring.

Chapter 6

Conclusion and Recommendations

6.1. Conclusion

Kvasir dataset integration with a potent deep learning algorithm and the in-click mobile application were effectively achieved in this project, enabling the classification of Gastro- intestinal diseases. The next uses Image Classification with Hltn, feature extraction with Vit where efficiency achieved is 91%. The last developed model was integrated into a mobile application where the users can upload gastrointestinal images and get the results quickly. This solution is useful to the healthcare providers and the patient in diagnosing gastrointestinal diseases in their initial stage hence improving management.

6.2. Recommendation

For future work, emphasis should be placed on refinement and fine-tuning as well as using other datasets, including more practical clinical images to remove doubt about the stability of the developed model. Additional functions, into the seat of application such as historical data analysis, personalized advises, /feedbacks from patients can help upscale the usage of the application. User information should be protected effectively from external threats due to powerful data privacy and security protocols. Inclusion of the system into healthcare frameworks and cloud solutions can contribute to the large-scale application that improves disease management outcomes for patients.

References

- [1] G. K. P. M. M. T. Subhashree Mohapatra, "Gastrointestinal abnormality detection and classification using empirical wavelet transform and deep convolutional neural network from endoscopic images," *Science Direct*, vol. 14, 2023.
- [2] M. S. Hossain, M. M. Rahman and M. F. U. M. H. M. A. H. Mahbubul Syeed, "DeepPoly: Deep Learning-Based Polyps Segmentation and Classification for Autonomous Colonoscopy Examination," *IEEE*, pp. 95889 - 95902, 2023.
- [3] K. T. S. H. T. H. Kyoka Yoshiok, "Deep-learning models in medical image analysis: Detection of esophagitis from the Kvasir Dataset," *arxiv*, 2023.
- [4] M. A. K. A. O.-Y. S. U. T. d. S. K. Javeria Naz, "Segmentation and Classification of Stomach Abnormalities Using Deep Learning," *Tech Science Press*, 2021.
- [5] İ. Korkmaz and F. Soygazi, "Gastrointestinal Image Classification Using Deep Learning Architectures via Transfer Learning," *IEEE Xplore*, 2024.
- [6] H. Ü. & H. F. Ayşe Ayyüce Demirbaş, "Spatial-attention ConvMixer architecture for classification and detection of gastrointestinal diseases using the Kvasir dataset," *Springer nature Link*, vol. 12, 2024.
- [7] R. M. 2. K. C. 3. T. R. M. 4. K. V. 5. T. E. Y. A M J MD Zubair Rahman 1, "Enhancing image- based diagnosis of gastrointestinal tract diseases through deep learning with EfficientNet and advanced data augmentation techniques," *PMC PubMed Central*, 2024.
- [8] Kvasir-Dataset link: <https://datasets.simula.no/kvasir/>