VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"JNANA SANGAMA", BELAGAVI -590 018.

2025 – 2026

A Project Report (Phase – II)
on

"Automated Ovarian Cancer Diagnosis Using CNN and Medical Imaging"

Submitted in partial fulfillment for the award of the degree of
BACHELOR OF ENGINEERING
in

CSE (Data Science)

Submitted By

Abrar-ul-haq	[1RN22CD004]
Adithya Sharma KS	[1RN22CD007]
C Keertan Kumar	[1RN22CD021]
Piyush Kumar	[1RN22CD060]

Under the guidance of

Dr. Sreevidya R C
Associate Professor, Dept. of CSE (DS)

Estd : 2001

DEPARTMENT OF CSE (DATA SCIENCE

RNS INSTITUTE OF TECHNOLOGY

Autonomous Institution affiliated to Visvesvaraya Technological University, Belagavi Approved By AICTE, New Delhi. Accredited by NAAC 'A'+ Grade Channasandra, Dr. Vishnuvardhan Road, Bengaluru - 560 098

RN SHETTY TRUST®

RNS INSTITUTE OF TECHNOLOGY

Autonomous Institution affiliated to Visvesvaraya Technological University, Belagavi Approved By AICTE, New Delhi. Accredited by NAAC 'A'+ Grade Channasandra, Dr. Vishnuvardhan Road, Bengaluru - 560 098

DEPARTMENT OF CSE (DATA SCIENCE)

Estd : 2001

CERTIFICATE

Certified that the project work (phase – II) entitled ***"Automated ovarian cancer diagnosis using CNN and medical imaging"*** carried out by Abrar-ul-haq **[1RN22CD004],** Adithya Sharma KS **[1RN22CD007],**C Keertan Kumar **[1RN22CD021],** Piyush Kumar **[1RN22CD060]** are bonafide students of **RNS Institute of Technology** in partial fulfilment for the award of **"BACHELOR OF ENGINEERING"** in **CSE (DATA SCIENCE)** as prescribed by **VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI** during the academic year **2025 – 26**. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the said degree.

Dr . Sreevidya R C	**Dr. Mohan H S**	**Dr. Ramesh Babu H S**
Guide & Associate Professor	**Professor & Head**	**Principal**
Dept. of CSE(DS), RNSIT	**Dept. of CSE(DS), RNSIT**	**RNSIT**

External Viva-Voce

Name of the Examiner	**Signature**
1. ………………………………..	………………..
2. ………………………………..	……………….

ACKNOWLEDGEMENT

We profoundly thank the **Management of RNS Institute of Technology** for their commitment to ensuring students have access to the best facilities and academic opportunities. Their constant support has made a major difference in the execution of this project.

We would like to express our sincere gratitude to **Dr. M K Venkatesha,** the **Director of RNSIT**, for his constant support, encouragement, and guidance throughout the duration of this project.

We are extremely grateful to **Dr. Ramesh Babu H S, Principal of RNSIT**, for providing the necessary resources and fostering an environment conducive to academic excellence. His leadership and vision have been a constant source of inspiration.

We would like to express our sincere appreciation to **Dr. Mohan H S**, **Head of the Department, CSE (Data Science)**, for his guidance and for ensuring that we had the proper direction in our academic endeavours. His encouragement throughout this project has been invaluable.

We would like to sincerely thank our project guide, **Dr. Sreevidya R C**, Associate Professor of the Department CSE (Data Science), for her invaluable guidance, continuous support, and expert advice throughout this project. Her encouragement and constructive feedback have played a significant role in the successful completion of this work.

We are deeply grateful to all the faculty members of the Department of CSE (Data Science) for their insights and encouragement, which have greatly helped in shaping the project.

We express our truthful thanks to **Dr. Mahantesh K,** Associate Professor & Project Coordinator, Dept. of CSE (Data Science) for his valuable support.

Finally, we take this opportunity to extend our earnest gratitude and respect to our parents, teaching & technical staff of the department, the library staff and all our friends, who have directly or indirectly supported us during the period of our project work.

Regards,

Abrar-ul-Haq	**[1RN22CD004]**
Adithya Sharma KS	**[1RN22CD007]**
C Keertan Kumar	**[1RN22CD021]**
Piyush Kumar	**[1RN22CD060]**

RN SHETTY TRUST®
RNS INSTITUTE OF TECHNOLOGY
Autonomous Institution affiliated to Visvesvaraya Technological University, Belagavi Approved By AICTE, New Delhi. Accredited by NAAC 'A'+ Grade Channasandra, Dr. Vishnuvardhan Road, Bengaluru - 560 098

Estd : 2001

DEPARTMENT OF CSE (DATA SCIENCE)

DECLARATION

I/We declare that this project report titled **Automated ovarian cancer diagnosis using CNN and medical imaging** submitted in partial fulfilment of the degree of **Bachelor of Engineering** in **CSE (Data Science)** is a record of original work carried out by me under the supervision of **Dr. Sreevidya R C**, and has not formed the basis for the award of any other degree, in this or any other Institution or University. In keeping with the ethical practice in reporting scientific information, due acknowledgements have been made wherever the findings of others have been cited.

Signature

Abrar-ul-Haq	**[1RN22CD004]**
Adithya Sharma K S	**[1RN22CD007]**
C Keertan Kumar	**[1RN22CD021]**
Piyush Kumar	**[1RN22CD060]**

Table of Contents

List of figures

Abstract

Ovarian cancer is one of the most fatal gynecological malignancies due to its asymptomatic progression and delayed diagnosis. Early detection and accurate staging of ovarian cancer are critical to improving patient survival rates and optimizing treatment planning. This project presents a machine learning-based approach for automated detection and staging of ovarian cancer using the publicly available TCGA-OV dataset.

The proposed methodology involves preprocessing clinical data and radiological imaging data (DICOM), followed by the implementation of two separate classification models: a Random Forest classifier for binary cancer detection using tabular clinical features, and a Convolutional Neural Network (CNN) model for multi-class cancer stage classification (Stage I–IV) using CT scan images. Extensive preprocessing, feature extraction, and model evaluation techniques were employed to ensure data consistency and predictive accuracy.

Experimental results demonstrate high accuracy in both cancer detection and stage classification, with the CNN model achieving an overall staging accuracy of 85%. The study highlights the potential of machine learning in augmenting clinical diagnostics by providing a fast, scalable, and accurate tool for ovarian cancer screening.

This work lays a foundation for future enhancements involving larger datasets, genomic integration, and real-time deployment in clinical settings.

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND OF OVARIAN CANCER

Ovarian cancer is one of the most severe and life-threatening gynecological malignancies affecting women worldwide. It originates in the ovaries, which are responsible for producing eggs and essential reproductive hormones. One of the major reasons ovarian cancer has a high mortality rate is its asymptomatic nature during the early stages. Most patients do not experience noticeable symptoms until the disease has progressed to an advanced stage, by which time effective treatment becomes significantly more challenging. According to clinical studies, a large percentage of ovarian cancer cases are diagnosed only at Stage III or Stage IV, resulting in reduced survival rates and limited therapeutic options.

From a clinical perspective, early diagnosis plays a crucial role in improving patient outcomes. When ovarian cancer is detected at an early stage, the five-year survival rate can be as high as 90%. However, due to the lack of reliable early screening techniques and subtle imaging characteristics, early detection remains a major challenge. Medical imaging modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and ultrasound are widely used to evaluate ovarian abnormalities, but accurate interpretation depends heavily on the expertise and experience of radiologists.

1.2 ROLE OF MEDICAL IMAGING IN OVARIAN CANCER DIAGNOSIS

Medical imaging forms the backbone of ovarian cancer diagnosis, staging, and treatment planning. Among various imaging techniques, CT scans are commonly used because they provide detailed cross-sectional images of the abdominal and pelvic regions. CT imaging helps in identifying tumor size, location, spread to nearby organs, lymph node involvement, and the presence of ascites or metastasis. These factors are critical for staging ovarian cancer and determining appropriate treatment strategies.

However, CT images consist of hundreds of slices per patient, making manual analysis extremely time-consuming. Additionally, ovarian tumors often exhibit subtle visual patterns that are difficult to distinguish from normal anatomical structures. Variations in contrast, scanner settings, patient anatomy, and slice orientation further complicate the diagnostic process. As a result, manual interpretation is prone to inter-observer variability and human error, particularly in busy clinical environments.

These challenges highlight the need for automated systems capable of assisting radiologists by analyzing CT images efficiently and consistently. Advances in artificial intelligence and deep learning have opened new possibilities for extracting meaningful patterns from medical images, enabling computer-aided diagnosis with high accuracy.

1.3 ARTIFICIAL INTELLIGENCE IN HEALTHCARE

Artificial Intelligence (AI) has rapidly transformed multiple domains, including healthcare, by enabling machines to learn from data and make intelligent decisions. In medical applications, AI systems are increasingly used for disease diagnosis, prognosis, treatment recommendation, and workflow optimization. Machine learning and deep learning techniques have demonstrated exceptional performance in analyzing complex and high-dimensional data such as medical images, electronic health records, and genomic information.

In recent years, deep learning models—especially **Convolutional Neural Networks (CNNs)**—have gained significant attention in medical image analysis. CNNs are specifically designed to process image data by learning hierarchical spatial features directly from pixel values. Unlike traditional image processing techniques that rely on handcrafted features, CNNs automatically learn discriminative features such as edges, textures, shapes, and complex anatomical patterns. This makes them particularly suitable for detecting tumors and abnormalities in medical images.

The integration of AI into radiology aims not to replace clinicians but to assist them by providing faster, more consistent, and data-driven insights. AI-based diagnostic tools can serve as second-opinion systems, reduce radiologist workload, and improve diagnostic accuracy, especially in resource-constrained healthcare settings.

1.4 MOTIVATION FOR AUTOMATED OVARIAN CANCER DETECTION

The motivation behind this project arises from the critical need for early and accurate ovarian cancer detection. Manual diagnosis using CT scans requires extensive expertise and time, and subtle tumor characteristics may be overlooked, especially during early stages. An automated system capable of analyzing CT images and identifying cancerous patterns can significantly enhance diagnostic efficiency and reliability.

Another key motivation is the increasing availability of large, publicly accessible medical datasets such as the **TCGA-OV (The Cancer Genome Atlas – Ovarian Cancer)** dataset. This dataset provides real-world clinical CT images in DICOM format along with associated metadata, enabling the development of realistic and clinically relevant machine learning models. Leveraging such datasets allows researchers to design AI systems that generalize well to real-world medical scenarios.

Furthermore, advancements in computational power and open-source deep learning frameworks such as PyTorch have made it feasible to train sophisticated CNN models even in academic environments. By combining these technological advancements with medical imaging, this project aims to develop an automated ovarian cancer detection system that is accurate, interpretable, and deployable.

1.5 PROBLEM STATEMENT

Despite significant progress in medical imaging and oncology, ovarian cancer diagnosis still faces several challenges. The absence of reliable early screening methods, subtle imaging features, and dependency on manual interpretation often result in delayed diagnosis. Radiologists must analyze large volumes of CT images, which increases cognitive load and the risk of oversight. Additionally, variability in interpretation among radiologists can lead to inconsistent diagnostic outcomes. Therefore, there is a pressing need for an automated, consistent, and scalable system that can analyze CT scan images and assist clinicians in detecting ovarian cancer. Such a system should be capable of learning complex spatial patterns from CT images, differentiating between cancerous and non-cancerous tissue, and providing explainable predictions to support clinical decision-making.

1.6 OBJECTIVES OF THE PROJECT

The primary objective of this project is to design and implement an automated ovarian cancer detection system using deep learning techniques and medical imaging data. The system aims to analyze CT scan and classify them as cancerous or non-cancerous with high accuracy.

The specific objectives of the project include:

- Preprocessing raw DICOM CT images to generate uniform, model-ready datasets.
- Developing a CNN-based model using the ResNet50 architecture for ovarian cancer detection.
- Applying transfer learning to improve performance with limited medical data.
- Incorporating Grad-CAM to provide visual explanations for model predictions.
- Deploying the trained model using a Flask-based web application for real-time inference.
- Evaluating the model using standard performance metrics such as accuracy and validation loss.

1.7 SCOPE OF THE PROJECT

The scope of this project is focused on automated ovarian cancer detection using CT scan images. The system is designed as a decision-support tool that can assist radiologists and clinicians by providing fast and consistent predictions. The project primarily addresses binary classification between cancerous and non-cancerous CT images, with stage classification explored as an extension.

While the system demonstrates strong performance on the TCGA-OV dataset, it is not intended to replace clinical diagnosis. Instead, it serves as a supportive AI tool that can enhance diagnostic workflows. Future improvements may include multi-stage cancer classification, integration of 3D CNNs for volumetric analysis, incorporation of clinical and genetic data, and deployment in hospital environments.

CHAPTER 2

LITERATURE SURVEY

2.1 INTRODUCTION TO LITERATURE SURVEY

1. [M. Radhakrishnan et al. IEEE Journal (2024)] - "Advancing Ovarian Cancer Diagnosis Through Deep Learning and Explainable AI: A Multiclassification Approach"
    - This study proposed a deep learning framework integrated with Explainable AI (XAI) for classifying ovarian cancer subtypes.
    - It achieved an impressive classification accuracy of 97.96% and provided interpretability by visually highlighting regions of importance in the input images.
    - The study did not explore generalizability across diverse datasets and real-time clinical environments.

2. [M. El-Khatib et al. IEEE Journal (2024)] - "New Trends in Ovarian Cancer Diagnosis Using Deep Learning: A Systematic Review"
    - This paper presented a comprehensive review of recent deep learning techniques used for ovarian cancer detection.
    - It emphasized the use of multiple neural networks and ensemble techniques to improve prediction accuracy.
    - The paper mainly aggregated previous findings and did not propose or validate a specific model implementation.

3. [Sunil Kumar Prabhakar et al. IEEE Journal (2020)] - "An Integrated Approach for Ovarian Cancer Classification with the Application of Stochastic Optimization"
    - The authors combined gene expression analysis with feature selection and optimization algorithms for classifying ovarian cancer using microarray data.
    - They showed that stochastic optimization improved the accuracy of gene-based classifiers.
    - The model was limited to genomic data and did not incorporate imaging modalities like CT or MRI.

2.2 TRADITIONAL METHODS FOR OVARIAN CANCER DIAGNOSIS

Traditionally, ovarian cancer diagnosis relies on a combination of clinical examination, blood tests such as CA-125 tumor markers, and imaging modalities including ultrasound, CT scans, and MRI. Among these, CT imaging is widely used for evaluating tumor spread and staging. Radiologists manually examine CT images to identify abnormal masses, ascites, lymph node involvement, and metastasis.

Although these methods are clinically effective, they suffer from several limitations. Manual interpretation of CT scans is highly dependent on the radiologist's expertise and experience. Subtle tumors or early-stage abnormalities may be overlooked, especially when image quality is poor or when radiologists are under heavy workload. Additionally, inter-observer variability often leads to inconsistent diagnostic outcomes. These limitations have encouraged researchers to explore automated and computer-aided diagnostic systems.

2.3 MACHINE LEARNING IN MEDICAL IMAGE ANALYSIS

Machine learning techniques have been widely applied to medical image analysis tasks such as disease classification, lesion detection, and segmentation. Early approaches relied on handcrafted feature extraction techniques, where features such as texture, shape, intensity histograms, and edge information were manually designed. These features were then fed into classifiers such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Random Forests.

While these traditional machine learning models showed moderate success, they extensive domain knowledge and manual feature engineering. Moreover, handcrafted features often failed to capture complex spatial patterns present in medical images. As a result, these approaches lacked robustness and generalization when applied to diverse datasets. The limitations of traditional machine learning paved the way for deep learning-based solutions.

2.4 DEEP LEARNING AND CONVOLUTIONAL NEURAL NETWORKS

Deep learning has revolutionized medical image analysis by enabling models to automatically learn hierarchical features directly from raw images. Convolutional Neural Networks (CNNs) are the most widely used deep learning architecture for image-based tasks. CNNs consist of convolutional layers, pooling layers, and fully connected layers that collectively learn low-level features such as edges and textures, as well as high-level semantic features such as shapes and spatial structures.

Several studies have demonstrated that CNNs outperform traditional machine learning techniques in medical imaging tasks. CNN-based models have been successfully applied to cancer detection in breast imaging, lung CT scans, brain MRI, and skin lesion analysis. These models exhibit superior accuracy, robustness, and scalability, making them suitable for large-scale medical applications.

In ovarian cancer research, CNNs have shown promise in analyzing CT and ultrasound images. However, the limited availability of labeled ovarian cancer datasets has restricted large-scale experimentation. This challenge has motivated the use of transfer learning techniques to leverage pretrained models.

2.5 TRANSFER LEARNING IN MACHINE LEARNING

Transfer learning is a widely adopted technique in medical image analysis, particularly when datasets are small. In transfer learning, a model pretrained on a large dataset such as ImageNet is fine-tuned on a target medical dataset. This allows the model to reuse low-level features learned from natural images while adapting higher-level layers to medical imaging tasks.

Architectures such as VGG16, ResNet50, DenseNet, and Inception have been commonly used in medical applications through transfer learning. Among these, ResNet50 is particularly effective due to its residual connections, which mitigate the vanishing gradient problem and allow deeper networks to converge efficiently.

Several studies have reported improved performance when using ResNet-based architectures for tumor detection and classification. Transfer learning reduces training time, improves generalization, and enables deep models to perform well even with limited medical data. This project adopts ResNet50 as the backbone architecture for ovarian cancer detection, inspired by its proven success in similar applications.

2.6 OVARIAN CANCER DETECTION USING DEEP LEARNING

Recent research has explored the use of deep learning for ovarian cancer diagnosis and staging. Radhakrishnan et al. (2024) proposed a CNN-based multiclass classification system integrated with explainable AI techniques for ovarian cancer subtype detection. Their model achieved high accuracy and demonstrated the importance of interpretability in clinical AI systems. However, the study was limited to a controlled dataset and did not address deployment challenges.

El-Khatib et al. (2024) conducted a systematic review of deep learning methods used for ovarian cancer diagnosis. The review highlighted the effectiveness of ensemble models and deep CNNs but also emphasized the lack of standardized datasets and real-time clinical validation. Most existing studies focused on experimental accuracy without addressing usability and deployment.

Other studies have explored ovarian cancer classification using genomic, proteomic, and clinical data, focusing on molecular biomarkers, gene expression patterns, and patient health records to support diagnosis and prognosis. While these approaches provide valuable insights into the biological mechanisms and risk stratification of ovarian cancer, they do not directly address imaging-based diagnosis, which is critical for early detection and clinical decision-making.

2.7 EXPLAINABLE AI IN MEDICAL IMAGING

One of the major challenges of deep learning models in healthcare is their black-box nature. Clinicians require transparency and interpretability to trust AI-driven predictions. Explainable Artificial Intelligence (XAI) techniques aim to address this issue by providing insights into model decision-making.

Grad-CAM (Gradient-weighted Class Activation Mapping) is a popular XAI technique used in CNN-based image classification. Grad-CAM generates heatmaps that highlight regions in the image that contribute most to the model's prediction. In medical imaging, these heatmaps help clinicians verify whether the model focuses on clinically relevant regions.

Several studies have emphasized that integrating explainability significantly improves clinical acceptance of AI systems. This project incorporates Grad-CAM to ensure that predictions are interpretable and medically meaningful.

2.8 CHALLENGES

- **Dataset Scarcity:** Ovarian cancer CT datasets are rare. TCGA-OV includes only cancerous cases thus additional non-cancerous data must be manually incorporated.
- **DICOM Variability:** Different scanners produce variations in pixel spacing, slice thickness, and reconstruction algorithms. This affects model learning and requires meticulous preprocessing.
- **Class Imbalance:** Cancerous images dominate datasets, while normal ovarian CT scans are limited, making training unbalanced classifiers difficult.
- **Model Interpretability:** Clinical acceptance demands explainable predictions. Deep CNNs are often treated as black boxes without visualization tools.
- **Patient Privacy and Data Security:** Handling medical imaging and clinical data requires strict adherence to data privacy regulations. Ensuring patient anonymity and secure processing of sensitive information is a critical challenge, especially when working with real-world healthcare datasets.

2.9 RESEARCH GAP AND MOTIVATION FOR THE PROPOSED WORK

The detailed review of existing literature on ovarian cancer diagnosis and medical image analysis reveals several critical research gaps that motivate the proposed work. Although numerous studies have demonstrated the effectiveness of deep learning techniques in cancer detection, most of the existing research focuses on cancers such as breast cancer, lung cancer, and brain tumors. Comparatively, ovarian cancer has received significantly less attention, primarily due to the limited availability of large, well-annotated imaging datasets. As a result, there is a lack of robust, generalized models specifically designed for ovarian cancer detection using CT imaging.

Another important gap identified in the literature is the limited focus on end-to-end system development. Many research works concentrate solely on improving classification accuracy under controlled experimental settings. These studies often overlook practical aspects such as data preprocessing from raw DICOM files, handling irrelevant CT slices, managing class imbalance, and deploying models for real-time usage. Consequently, while high accuracy values are reported in research papers, the proposed methods are rarely translated into usable systems that can assist clinicians in real-world scenarios.

Furthermore, most existing ovarian cancer studies rely on either small, curated datasets or single-institution data, which restricts model generalization. Variability in CT scanners, imaging protocols, slice thickness, and contrast levels is a major challenge in medical imaging. However, many studies do not explicitly address these variations during preprocessing and model training. This creates a gap in developing models that are robust enough to handle heterogeneous clinical data. The proposed project attempts to address this issue by using the TCGA-OV dataset, which represents real-world clinical diversity, and by applying systematic preprocessing and filtering techniques.

CHAPTER 3

SYSTEM ANALYSIS AND DATASET DESCRIPTION

3.1 OVERVIEW OF THE SYSTEM

The proposed system is designed as an end-to-end automated framework for ovarian cancer detection and analysis using deep learning techniques and medical imaging data. The system takes CT scan images in DICOM format as input, processes them through multiple preprocessing and filtering stages, and applies a Convolutional Neural Network (CNN) to classify the images as cancerous or non-cancerous. Additionally, the system explores cancer stage prediction as an extension using labeled clinical data. The final output is presented to the user through a Flask-based web application, enabling real-time interaction and visualization.

The system is developed with a strong emphasis on accuracy, interpretability, and usability. By integrating Grad-CAM visualization techniques, the system ensures that predictions are explainable, which is essential for clinical adoption. The modular design of the system allows each stage—data preparation, model training, prediction, and deployment—to operate independently, facilitating easier debugging, enhancement, and scalability.

3.2 DATASET SOURCE: TCGA-OV

The primary dataset used in this project is the **TCGA-OV (The Cancer Genome Atlas – Ovarian Cancer)** dataset obtained from **The Cancer Imaging Archive (TCIA)**. This dataset contains CT scan images of ovarian cancer patients in DICOM format along with associated clinical metadata. The TCGA-OV dataset is widely used in academic research due to its authenticity, standardized format, and availability of clinical annotations.

The CT images in the dataset represent real-world clinical conditions and include variability in scanner types, imaging protocols, slice thickness, and resolution. Such variability introduces challenges but also ensures that the trained model generalizes well to diverse clinical scenarios Since the TCGA-OV dataset primarily consists of cancerous cases, an additional non-cancerous CT dataset was incorporated to enable binary classification.

Fig 3.1 Cancerous Dataset

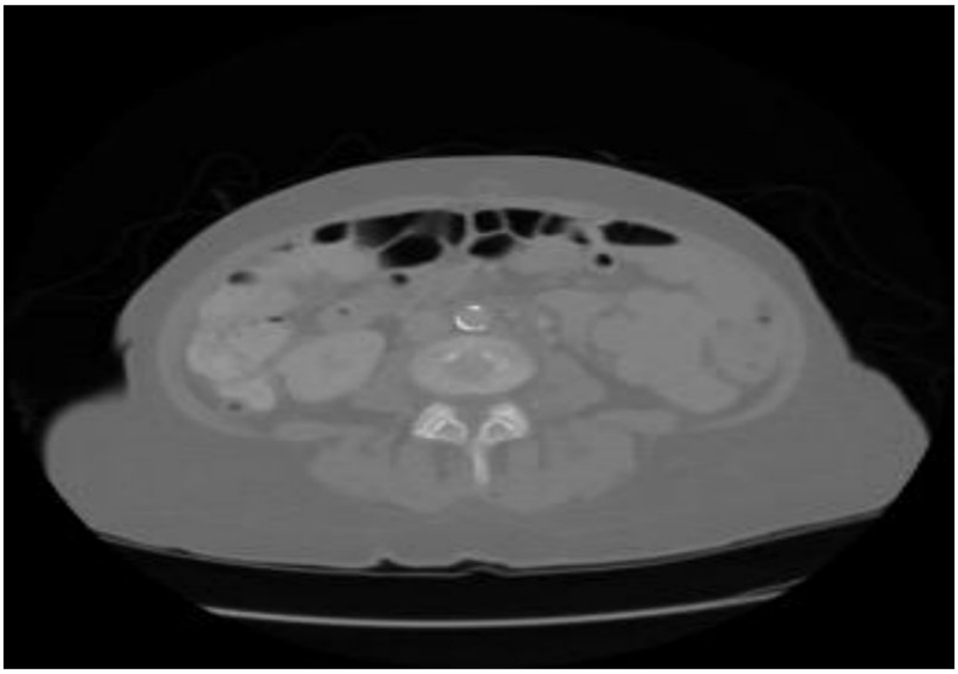

Fig 3.2 Cancerous Image

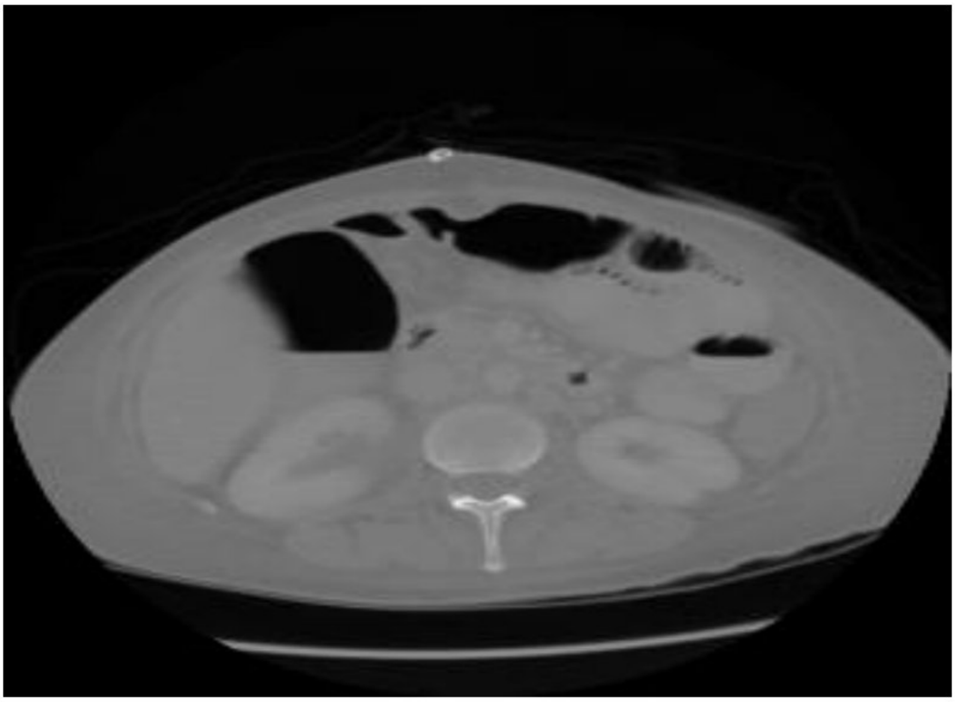

Fig 3.3 Cancerous Image

3.3 NON-CANCEROUS DATASET INTEGRATION

To enable effective binary classification between cancerous and non-cancerous images, normal abdominal CT images were collected from publicly available medical imaging repositories. These images do not contain ovarian cancer and serve as negative samples during training. The non-cancerous dataset underwent the same preprocessing pipeline as the cancerous dataset to ensure consistency in image quality, resolution, and intensity distribution.

The inclusion of non-cancerous data is critical for reducing bias and improving the robustness of the model. Without such data, the model would be unable to learn the distinguishing characteristics between healthy and abnormal ovarian tissue. Care was taken to ensure that the non-cancerous images were anatomically comparable to the ovarian CT scans used in the cancerous dataset.

Fig 3.4 Non-Cancerous Dataset

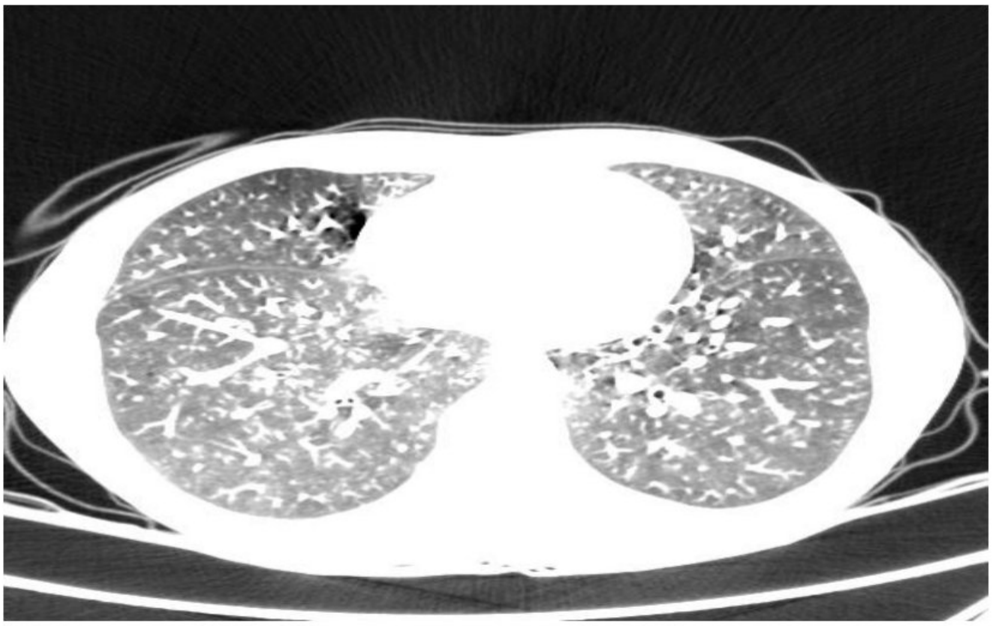

Fig 3.5 Non-Cancerous Image

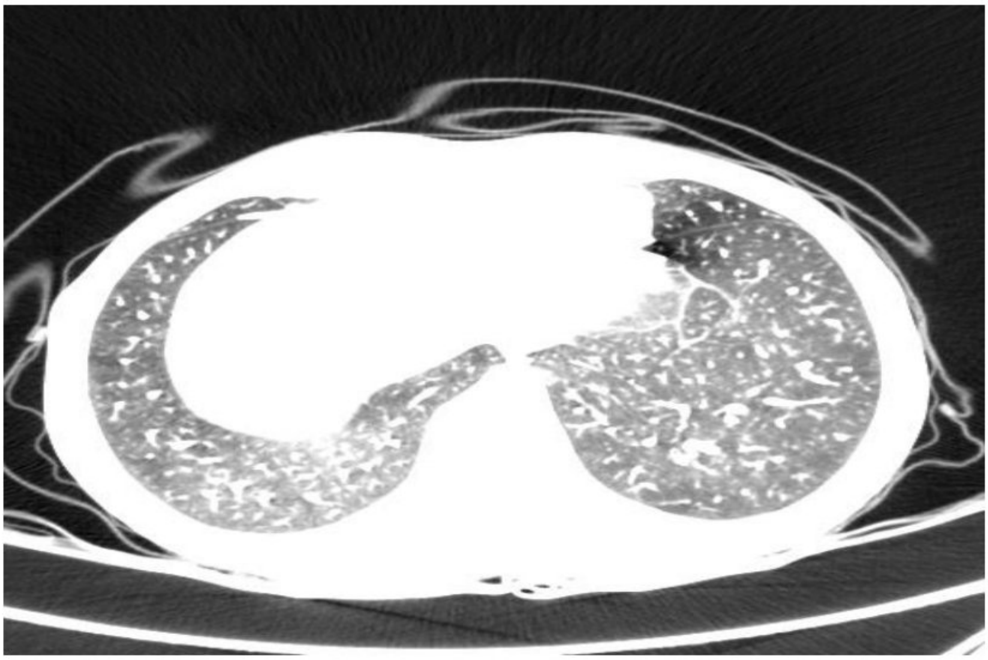

Fig 3.6 Non-Cancerous Image

3.4 DATA CHARACTERISTICS AND CHALLENGES

Medical imaging datasets present unique challenges compared to natural image datasets. CT scan images are grayscale and contain subtle intensity variations that represent tissue density differences. Additionally, each patient scan contains hundreds of slices, many of which do not include the ovarian region. Including such irrelevant slices can introduce noise and degrade model performance.

Another challenge lies in the DICOM format itself. DICOM images store metadata such as pixel spacing, slice thickness, and orientation, which vary across scanners. Inconsistent metadata can lead to distorted images if not handled correctly. Furthermore, the dataset exhibits class imbalance, particularly in stage-wise classification, as early-stage ovarian cancer cases are significantly fewer than advanced-stage cases.

3.5 DATA FILTERING STRATEGY

Initially, approximately 10,000 CT image slices were extracted from the combined datasets after Converting DICOM files to image format. However, not all slices were suitable for model training. A multi-level filtering strategy was applied to reduce the dataset to high-quality, relevant samples.

First, slices that did not contain the pelvic or ovarian region were removed. These included images from the chest, upper abdomen, and lower limbs. Second, low-quality slices affected by motion blur, poor contrast, or imaging artifacts were discarded. Third, for stage prediction, only images belonging to patients with valid clinical stage labels were retained. This filtering process significantly reduced the dataset size but improved overall data quality and relevance.

3.6 DATA REDUCTION AND AUGMENTAION

After filtering, the dataset size was reduced to a manageable subset suitable for training on limited hardware resources. To address class imbalance and enhance model generalization, data augmentation techniques were applied to the filtered dataset. Augmentation techniques included rotation, flipping, zooming, brightness adjustment, and noise addition.

Augmentation increases the effective size of the dataset without collecting new images, allowing the model to learn invariant features. In the case of stage prediction, augmentation was particularly important to balance the number of samples across different cancer stages. This strategy enabled the model to train effectively despite the limited availability of original stage-wise images.

After the completion of preprocessing and initial filtering, the combined dataset obtained from TCGA-OV and non-cancerous sources contained nearly 10,000 CT image slices. Although this dataset was sufficiently large in terms of quantity, it was not optimal for direct use in deep learning training. A significant portion of the images contained redundant information, repeated anatomical views, or minimal diagnostic relevance. Training a deep CNN such as ResNet50 on the full dataset would not only increase computational cost but could also negatively affect learning due to excessive redundancy.

To address this issue, a structured data reduction strategy was implemented. The reduction process focused on selecting representative CT slices that clearly displayed the pelvic region and ovarian anatomy. Slices that were visually similar, contained overlapping information, or did not contribute new diagnostic features were removed. This approach ensured that the retained dataset was compact yet informative, enabling the model to learn discriminative features more efficiently.

Another important reason for data reduction was the limitation of computational resources. Since the model was trained on a CPU-based environment without high-end GPUs, reducing dataset size helped maintain feasible training time and stable memory usage. This allowed multiple training iterations, hyperparameter tuning, and experimentation without excessive delays.

After reduction, attention was given to class balance. The dataset was reorganized to maintain a reasonable balance between cancerous and non-cancerous images. For stage-wise classification, data imbalance was more prominent, as early-stage ovarian cancer cases were significantly fewer than advanced-stage cases. To overcome this, data augmentation techniques were applied extensively.

Data augmentation artificially increases dataset diversity by applying transformations such as rotation, flipping, zooming, brightness variation, contrast adjustment, and noise addition. These transformations simulate real-world variations in CT imaging while preserving underlying medical meaning. Augmentation was applied dynamically during training so that the model encountered new variations in each epoch. This strategy improved generalization, reduced overfitting, and enabled effective learning even with limited original data.

3.7 SYSTEM ARCHITECTURE OVERVIEW

The system architecture consists of four major components: data preprocessing, model training, prediction with explainability, and deployment. Preprocessing modules handle DICOM conversion, normalization, resizing, and dataset organization. The model training module utilizes a pretrained ResNet50 architecture fine-tuned on the ovarian CT dataset.

The prediction module loads the trained model and applies Grad-CAM to generate heatmaps that highlight regions influencing predictions. Finally, the deployment module integrates the trained model and visualization outputs into a Flask web application. This layered architecture ensures modularity, maintainability, and ease of future enhancements.

The system architecture of the proposed ovarian cancer detection framework is designed in a modular and layered manner to ensure flexibility, scalability, and maintainability. Each module performs a specific function while seamlessly integrating with the others. This modular architecture simplifies debugging, testing, and future enhancements.

The first layer of the architecture is the data preprocessing layer, which handles DICOM image loading, normalization, resizing, and dataset organization. This layer ensures that raw medical images are transformed into a consistent and model-compatible format. By isolating preprocessing operations, the system can easily adapt to new datasets or imaging modalities in the future.

The second layer is the **model training layer**, which consists of the ResNet50-based CNN. This layer is responsible for feature extraction and classification. The CNN processes input images through multiple convolutional and residual blocks, learning hierarchical representations of ovarian tissue and tumour-related patterns.

The architecture allows transfer learning, enabling the reuse of pretrained features while fine-tuning task-specific layers. The third layer is the **prediction and explainability layer**, where trained models generate predictions for unseen CT images. This layer integrates Grad-CAM to produce heatmaps highlighting regions that influenced the model's decision. Explainability is treated as a core architectural component rather than an optional add-on.

The final layer is the **deployment layer**, implemented using a Flask web application. This layer provides a user-friendly interface for uploading CT images, triggering predictions, and displaying results. The layered architecture ensures smooth data flow from raw input to final output.

3.8 HARDWARE AND SOFTWARE REQURIMENTS

The system was developed and tested on a standard computing environment without the use of specialized GPUs. The hardware configuration includes a multi-core CPU, sufficient RAM, and local storage for dataset management. This demonstrates that the system is accessible and reproducible in typical academic settings.

The software stack includes Python as the primary programming language, PyTorch for deep learning, OpenCV and PIL for image processing, and Flask for web deployment. Additional libraries such as NumPy, pandas, and scikit-learn were used for data handling and evaluation.

The proposed system was developed with accessibility and reproducibility in mind, ensuring that it can operate on standard academic computing environments. The hardware requirements include a multi-core CPU, sufficient RAM for handling image datasets, and adequate storage for DICOM and processed image files. The system does not rely on specialized GPU hardware, making it suitable for institutions with limited computational resources.

From a software perspective, Python was chosen as the primary programming language due to its extensive support for machine learning and medical image processing. The PyTorch framework was used for deep learning model development, providing flexibility in model customization and training. Additional libraries such as NumPy and pandas were used for data handling, while OpenCV and PIL were employed for image preprocessing.

The Flask framework was used to develop the web-based deployment interface, allowing seamless interaction between the trained model and the end user. The use of open-source tools ensures cost-effectiveness and encourages reproducibility. Overall, the chosen hardware and software configuration supports efficient development, experimentation, and deployment of the system.

3.9 FEASIBILITY ANALYSIS

The proposed system is technically, economically, and operationally feasible. From a technical standpoint, the use of established deep learning frameworks and publicly available datasets ensures reliability. Economically, the system requires minimal infrastructure and can be executed on standard hardware. Operationally, the Flask-based interface enables easy interaction and deployment, making the system suitable for both academic and clinical demonstration purposes.

The feasibility of the proposed ovarian cancer detection system was evaluated across three major dimensions: technical feasibility, economic feasibility, and operational feasibility. From a technical standpoint, the use of proven deep learning architectures, established preprocessing techniques, and widely adopted frameworks ensures reliability and robustness. The system leverages publicly available datasets and open-source tools, minimizing dependency on proprietary resources.

Economic feasibility is achieved by designing the system to operate on standard computing hardware without the need for expensive GPUs or cloud infrastructure. This makes the solution affordable and suitable for academic institutions and small healthcare setups. The use of free and open-source software further reduces development and deployment costs.

CHAPTER 4

PROPOSED METHODOLOGY

4.1 INTODUCTION TO PROPOSED METHODOLOGY

The proposed methodology aims to develop a comprehensive and intelligent system for the automated detection of ovarian cancer using CT scan images and deep learning techniques. The methodology is designed to address the major challenges identified in the literature, such as variability in medical imaging data, limited availability of labeled datasets, lack of explainability in deep learning models, and absence of deployable solutions. Rather than focusing solely on classification accuracy, the methodology emphasizes data reliability, robustness of the model, interpretability of predictions, and real-time usability.

This methodology follows a structured, modular approach that begins with raw DICOM image acquisition and ends with an interactive web-based prediction system. Each stage of the pipeline—data preprocessing, filtering, augmentation, model training, prediction, and deployment—has been carefully designed based on the practical constraints of medical imaging data and the actual implementation carried out in this project. By combining transfer learning, CNN-based feature extraction, and explainable AI, the proposed methodology ensures that the system is not only accurate but also clinically meaningful.

- **Dataset Acquisition:** The system uses the TCGA-OV dataset consisting of cancerous ovarian CT scan DICOM images, along with an additional non-cancerous dataset collected from open-source abdominal CT repositories.
- **DICOM Preprocessing:** All DICOM images are converted into PNG format using prepare_dataset.py, where pixel intensity normalization, resizing, and contrast adjustments are applied.
- **Non-Cancerous Dataset Preparation:** The script prepare_noncanceroous.py processes normal CT scan images using the same preprocessing steps as cancerous images. This guarantees that both datasets maintain consistent quality and resolution for effective training.

- **Dataset Splitting:** The combined dataset is divided into training, validation, and test subsets using an 70:15:15 ratio. This ensures balanced learning and accurate evaluation of model generalization.
- **Model Architecture (ResNet50):** A pretrained ResNet50 model is fine-tuned using the ovarian CT dataset by replacing its fully connected layer with a custom binary classifier.
- **Model Training:** The training script train_resnet50.py trains the CNN using augmented images to improve robustness. The optimization uses Adam, and the best-performing model checkpoint is saved for prediction.
- **Grad-CAM Explainability:** Grad-CAM is integrated into predict_image.py to generate heatmaps showing the specific regions influencing model predictions.
- **Flask Deployment:** A lightweight Flask application (app_flask5.py) enables users to upload images directly through a browser interface. The system performs prediction, generates Grad-CAM heatmaps, and returns results in real time.
- Overall, the proposed method establishes a complete and efficient pipeline—from raw DICOM preprocessing to model training, explainability, and Flask deployment—ensuring reliable ovarian cancer detection from CT images.

4.2 OVERALL WORKFLOW OF THE SYSTEM

The overall workflow of the proposed system represents a sequential and interconnected process that transforms raw CT scan images into meaningful diagnostic outputs. The workflow begins with the acquisition of CT images in DICOM format from the TCGA-OV dataset and additional non-cancerous sources. These images are initially raw and unstructured, containing variations in orientation, intensity, and resolution.

Once acquired, the images undergo preprocessing and filtering to remove irrelevant and low-quality slices. The processed dataset is then divided into training, validation, and testing sets. The training data is used to fine-tune a pretrained ResNet50 model, enabling it to learn ovarian cancer–specific features. After training, the model is evaluated on unseen data to ensure generalization. During inference, the trained model predicts the presence or absence of cancer and generates Grad-CAM heatmaps that visually explain the prediction.

Finally, the complete workflow is deployed through a Flask-based web application, enabling users to upload CT images and receive diagnostic results interactively. This workflow ensures smooth data flow between modules, reduces redundancy, and enables systematic debugging and enhancement of individual components.

4.3 DATASET ACQUISITION AND SOURCES

The success of any deep learning system heavily depends on the quality and relevance of the dataset used. In this project, CT scan images were collected primarily from the TCGA-OV dataset hosted by The Cancer Imaging Archive. This dataset contains real clinical CT scans of ovarian cancer patients, making it highly suitable for developing and evaluating a diagnostic system. The images are provided in DICOM format, which is the standard format used in medical imaging systems.

Each CT scan in the dataset consists of multiple axial slices capturing the abdominal and pelvic regions. These scans reflect real-world imaging conditions, including variations in scanner models, acquisition parameters, and patient anatomy. Such variability poses challenges but also ensures that the trained model learns generalized features rather than overfitting to a specific imaging protocol.

Since TCGA-OV primarily contains cancerous cases, additional non-cancerous abdominal CT images were incorporated to create a balanced dataset. These images serve as negative samples, allowing the model to learn the differences between normal and abnormal ovarian tissue.

4.4 DICOM IMAGE PROCESSING

DICOM images contain both pixel data and metadata, which must be handled carefully before being used in deep learning models. Directly feeding raw DICOM images into a CNN is impractical due to inconsistent pixel intensity scales and varying image dimensions. Therefore, a robust preprocessing pipeline was implemented using custom Python scripts. The preprocessing begins with reading DICOM files using specialized libraries and extracting pixel arrays.

Window-level and window-width adjustments are applied to enhance soft-tissue contrast, which is essential for identifying ovarian abnormalities. Pixel intensities are normalized to a consistent range to reduce scanner-dependent variations. The images are then resized to a fixed resolution of 224×224 pixels to meet the input requirements of the ResNet50 architecture.

Noise reduction techniques are applied to suppress artifacts caused by imaging equipment or patient movement. The final preprocessed images are converted into PNG format and stored in an organized directory structure, making them compatible with PyTorch data loaders.

4.5 FILTERING OF CT SCAN SLICES

CT scans consist of a large number of slices, many of which are irrelevant for ovarian cancer analysis. Including such slices can introduce noise and degrade model performance. Therefore, an extensive filtering process was employed to retain only clinically meaningful images.

First, slices that did not belong to the pelvic region were removed. These included slices capturing the chest, upper abdomen, and lower extremities. Second, slices with poor visual quality, such as blurred images or slices with extremely low contrast, were discarded. Third, for stage classification, only slices associated with patients having valid and clearly defined clinical stage labels were retained.

This filtering process significantly reduced dataset size while improving data relevance. Although filtering reduced the number of available images, it ensured that the model trained only on meaningful data, leading to better feature learning and improved accuracy.

4.6 DATASET REDUCTION STRATEGY

After preprocessing and filtering, the dataset initially contained nearly 10,000 CT image slices. Training a deep CNN on this large dataset using limited computational resources posed challenges in terms of training time and memory usage. Additionally, many slices contained redundant information that contributed minimally to learning.

To address these issues, a dataset reduction strategy was applied. A representative subset of images was selected based on image clarity, anatomical relevance, and class balance. This reduction enabled faster experimentation, efficient hyperparameter tuning, and stable training. Importantly, the reduced dataset retained sufficient diversity to allow the model to learn robust features. Dataset reduction also played a key role in ensuring that the system could be trained and demonstrated on standard academic hardware, making the project reproducible and practical.

4.7 DATA AUGMENTATION TECHNIQUES

Data augmentation is an essential technique for improving model generalization, especially when working with limited medical datasets. In this project, augmentation was applied after dataset reduction to artificially increase data diversity. Augmentation techniques included rotation, flipping, scaling, brightness adjustment, and noise addition.

These transformations simulate real-world variations in patient positioning, scanner calibration, and imaging conditions. Augmentation ensures that the model does not memorize specific patterns but instead learns invariant and generalizable features. For stage classification, augmentation was particularly important to balance the number of samples across different cancer stages.

By applying augmentation dynamically during training, the model was exposed to a wider range of variations, reducing overfitting and improving robustness.

4.8 CNN BASED MODEL SELECTION

Convolutional Neural Networks (CNNs) form the core of the proposed ovarian cancer detection system due to their exceptional ability to process and analyze image-based data. Unlike traditional machine learning models that rely on manually engineered features, CNNs automatically learn hierarchical feature representations directly from raw pixel values. This capability is particularly important in medical imaging, where critical diagnostic information is often embedded in subtle texture variations, shape irregularities, and spatial patterns that are difficult to capture using handcrafted features.

A CNN architecture is composed of multiple layers, each serving a specific purpose in the feature extraction and classification process. The convolutional layers apply learnable filters to input images to detect local patterns such as edges, contours, and texture gradients. In the context of CT scan images, these filters help identify tissue boundaries, density variations, and abnormal growth patterns associated with ovarian tumors. As the network depth increases, the convolutional layers learn increasingly abstract and high-level features, such as tumor shape, structural deformation, and abnormal mass distribution within the pelvic region.

Pooling layers are used in conjunction with convolutional layers to reduce spatial dimensions and computational complexity while retaining essential information. Max pooling operations help the model become invariant to small translations and noise, which is critical when dealing with CT scans obtained under varying patient positions and scanner settings. Among various CNN architectures available, ResNet50 was selected as the backbone model for this project due to its depth and residual learning mechanism.

Traditional deep CNNs often suffer from the vanishing gradient problem, where gradients diminish as they propagate through deeper layers, making training difficult. ResNet addresses this issue through residual connections, which allow the network to learn identity mappings and propagate gradients more effectively.

4.9 TRANSFER LEARNING STRATEGY

Transfer learning plays a crucial role in this project due to the limited availability of labeled medical images. By initializing the ResNet50 model with pretrained ImageNet weights, the model benefits from previously learned visual features. These features serve as a strong foundation for learning medical image characteristics.

The final classification layers of the network were replaced with custom layers tailored for ovarian cancer detection and stage classification. During training, earlier layers were partially frozen to retain generic feature representations, while later layers were fine-tuned to learn task-specific patterns. This approach significantly reduced training time and improved convergence stability.

4.10 MODEL TRAINING PROCEDURE

The model training process was implemented using the PyTorch framework. The filtered and augmented dataset was divided into training, validation, and testing subsets. The training data was used to update model weights, while validation data was used to monitor performance and tune hyperparameters.

The Adam optimizer was employed due to its adaptive learning rate mechanism, which accelerates convergence. Binary Cross-Entropy loss was used for cancer detection tasks, while categorical loss functions were applied for stage classification. Training was conducted over multiple epochs, and model checkpoints were saved based on validation performance.

4.11 PERFORMANCE MONITORING DURING TRAINING

Continuous performance monitoring was a critical component of the training process. Metrics such as training accuracy, validation accuracy, and loss values were recorded for each epoch. These metrics provided insights into learning behavior and helped identify overfitting or underfitting.

Validation performance was used to determine the optimal stopping point for training. Early stopping techniques were applied to prevent unnecessary training once performance stabilized. Training curves were analyzed to fine-tune hyperparameters and improve model stability. This systematic monitoring ensured that the final model achieved optimal performance while maintaining generalization capability.

In addition to quantitative metrics, qualitative analysis of model predictions was also carried out during training. Sample validation images were periodically evaluated to observe how confidently and consistently the model classified different stages. This helped in understanding misclassification patterns and class-wise performance, especially in cases where stage boundaries were visually subtle. Confusion matrices and class-wise accuracy trends were reviewed to ensure balanced learning across all categories. Such comprehensive monitoring not only improved interpretability of the model's behavior but also increased confidence in its robustness and reliability for real-world clinical application.

4.12 TRAINING STRATEGY AND HYPERPARAMETER SELECTION

The training strategy adopted in this project was carefully designed to ensure stable learning, effective generalization, and efficient utilization of limited computational resources. Since medical imaging datasets are complex and often limited in size, improper training strategies can easily lead to overfitting or unstable convergence. Therefore, a structured training pipeline was implemented using the PyTorch framework.

The filtered and augmented dataset was divided into three subsets: training, validation, and testing. Approximately 70% of the dataset was used for training the model, 15% for validation, and the remaining 15% for testing. This split ensured that the model was trained on sufficient data while also being evaluated on unseen samples to measure generalization capability. Data augmentation was applied only to the training set to prevent data leakage into validation and testing phases.

Hyperparameters such as batch size, learning rate, number of epochs, and dropout rate were tuned through experimental evaluation. Training was conducted for multiple epochs until validation performance stabilized. Early stopping was applied to prevent overfitting when validation accuracy stopped improving. Model checkpoints were saved based on the best validation performance, ensuring that the optimal model was used for inference and deployment.

Binary Model

Hyperparameter	Value
Optimizer	Adam
Learning Rate	0.0001
Batch Size	16
Loss Function	Binary Cross Entropy (BCEWithLogits → Sigmoid)
Epochs	20 (early stopped at ~8)
Dropout	0.3
Weight Initialization	ResNet-50 pretrained backbone or random

Fig. 4.1 Hyper-parameters

4.13 STAGE CLASSIFICATION METHODOLOGY

In addition to binary cancer detection, the project explores stage classification as an extended objective. Ovarian cancer staging plays a critical role in determining treatment strategies and prognosis. However, stage-wise classification presents additional challenges due to severe class imbalance and limited availability of labeled data.

Stage labels were obtained from the clinical metadata file associated with the TCGA-OV dataset. Only patients with clearly defined clinical stages were considered for this task. CT image slices corresponding to these patients were filtered to retain pelvic-region images relevant to ovarian analysis. Due to the limited number of samples in early stages, direct training without preprocessing would result in biased predictions.

To address this issue, extensive data augmentation was applied to balance stage-wise samples. After augmentation, the dataset was structured into three primary stages for classification: Stage I, Stage II, and Stage III. The same CNN backbone (ResNet50) was used for stage classification, with the final classification layer modified to output multiple classes using a Softmax activation function. The stage classification model was trained using categorical cross-entropy loss. During training, performance metrics such as class-wise accuracy and confusion matrices were analyzed to ensure that the model did not bias toward majority stages.

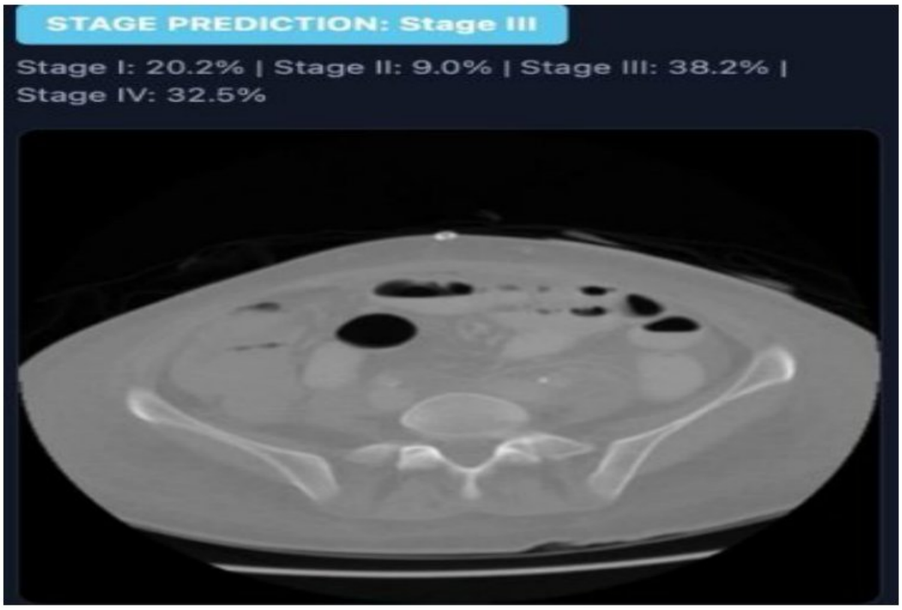

Fig. 4.2 Ovarian Cancer Stage Prediction

4.14 EXPLAINABILITY USING GRAD-CAM

Explainability is a critical requirement in medical AI systems, as clinicians must understand the reasoning behind automated predictions before trusting them. Deep learning models are often criticized for their black-box nature, which limits their adoption in clinical environments. To address this concern, the proposed system integrates **Gradient-weighted Class Activation Mapping (Grad-CAM)** as an explainability mechanism.

Grad-CAM works by computing the gradient of the predicted class score with respect to the feature maps of the final convolutional layer. These gradients indicate how important each spatial location is for the prediction. By weighting the feature maps using these gradients and projecting them back onto the input image, Grad-CAM produces a heatmap highlighting regions that influenced the model's decision.

In this project, Grad-CAM is applied during the prediction stage to generate visual explanations for both cancer detection and stage classification. For cancerous images, the heatmaps typically highlight irregular masses, dense tissue regions, or abnormal structural patterns in the pelvic area. For non-cancerous images, minimal activation is observed, indicating correct model behavior. Overall, the integration of Grad-CAM into the proposed ovarian cancer detection system plays a crucial role in bridging the gap between deep learning predictions and clinical interpretability. By visually highlighting the regions that influence the model's decisions, Grad-CAM transforms the CNN from a black-box model into a transparent and explainable system. This not only improves confidence in the model's predictions but also enables clinicians to cross-verify whether the highlighted regions correspond to medically relevant tumor areas.

The use of Grad-CAM therefore enhances trust, supports clinical decision-making, and ensures that the proposed system aligns with the ethical and practical requirements of medical AI applications. Consequently, explainability through Grad-CAM significantly strengthens the reliability and real-world applicability of the proposed methodology.

4.15 LOCAL WEB – BASED IMPLEMENTATION (FLASK)

The local web-based implementation of the proposed ovarian cancer detection system serves as an effective proof-of-concept that demonstrates the practical usability of the developed deep learning model. By integrating the trained CNN and Grad-CAM explainability into a Flask-based web interface, the system allows users to upload CT scan images and obtain predictions in an interactive and user-friendly manner. Although the implementation is limited to a local environment, it successfully validates the end-to-end workflow—from image input and preprocessing to prediction and visualization.

The Flask application provides a simple interface where users can upload CT scan images. Once an image is uploaded, it is passed to the backend for preprocessing, prediction, and Grad-CAM visualization. The trained model loads the saved weights and performs inference on the input image. The prediction result and corresponding heatmap are then returned to the user through the web interface.

This local web-based setup provides a strong foundation for future enhancements, such as cloud deployment or hospital-level integration, while clearly showcasing the feasibility, functionality, and real-time capability of the proposed system.

4.16 FLOW DIAGRAM

The flow diagram represents the complete operational pipeline of the proposed ovarian cancer detection system. It begins with the acquisition of CT scan images in DICOM format, followed by preprocessing steps such as normalization, resizing, and filtering. The processed images are then used to train the CNN model using a structured training strategy.

During inference, the trained model predicts cancer presence or stage and simultaneously generates Grad-CAM heatmaps to explain predictions. Finally, the results are displayed to the user through the Flask web interface. The flow diagram provides a clear visual overview of data movement across different system components.

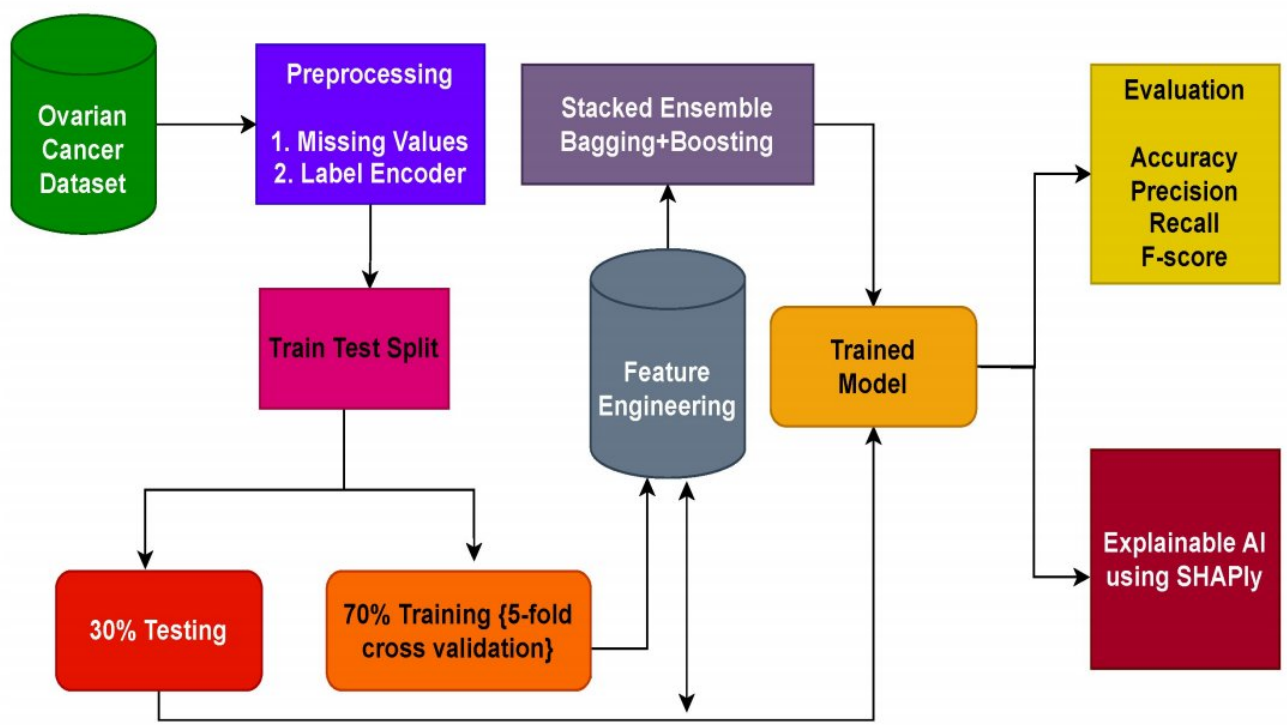

Fig. 4.3 Flow cycle of Ovarian Cancer Detection

1. **Dataset Collection (TCGA-OV & Non-Cancerous Data):** The process begins by collecting cancerous CT images from the TCGA-OV dataset and non-cancerous abdominal CT images from open-source repositories.
2. **Preprocessing of Images:** All images undergo preprocessing steps such as DICOM conversion, contrast normalization, resizing, and noise reduction.
3. **Model Training Using Processed Dataset:** The cleaned dataset is used to train the CNN-based ResNet50 model, which learns to differentiate between cancerous and healthy ovarian tissue.
4. **Grad-CAM Visualization for Interpretability:** Once the model is trained, Grad-CAM is applied to generate heatmaps that highlight the regions the network focuses on during prediction.
5. **Demonstration through Flask Application:** The final trained model and Grad-CAM module are integrated into a Flask-based application for real-time usage.

4.17 CURRENT PROGRESS

- Significant progress has been achieved in the implementation and validation of the ovarian cancer detection system. The dataset preparation scripts have successfully processed thousands of CT image slices, extracting clean, uniform-quality samples from the TCGA-OV dataset.

- Noise removal, normalization, and image resizing have been completed for both cancerous and non-cancerous sets. The system also includes a label verification mechanism through check_labels.py, ensuring that images are correctly mapped to their corresponding patient IDs and cancer statuses.

- The ResNet50 model has been fully trained using the processed dataset, achieving stable convergence and strong generalization. Multiple training cycles were executed, and the model demonstrated increasing accuracy with each epoch. Hyperparameters were fine-tuned to reduce overfitting, including dropout usage, learning-rate scheduling, and data augmentation.

- The Grad-CAM visualization has been successfully integrated into the prediction script. For several sample CT scans, the system correctly highlighted abnormal tissue regions, validating that the model is learning meaningful features rather than random noise. This enhances trust and provides a crucial layer of explainability.

- On the deployment side, the Flask application is fully operational. It accepts user uploads, passes input images through the prediction pipeline, and visually displays both classification output and Grad-CAM heatmaps. The entire system functions smoothly on CPU execution, making it accessible on standard laptop hardware.

CHAPTER 5

IMPLEMENTATION AND EXPERIMENTAL RESULTS

5.1 INTRODUCTION

This chapter focuses on the practical realization of the proposed automated ovarian cancer detection system described in the previous chapters. While Chapter 4 detailed the theoretical methodology and system design, the objective of this chapter is to explain how the proposed ideas were actually implemented using real datasets, programming frameworks, and computational resources. The implementation demonstrates the feasibility of translating deep learning concepts into a working system capable of analyzing medical images and generating meaningful diagnostic outputs.

The implementation phase involved converting raw medical imaging data into a structured and usable form, developing and training a convolutional neural network model, evaluating its performance using standard metrics, and integrating the trained model into a local web-based application. Each component of the system was implemented using Python and open-source libraries to ensure transparency, reproducibility, and ease of experimentation. Emphasis was placed on maintaining consistency between the proposed methodology and the actual execution while addressing practical challenges encountered during development.

This chapter also presents the experimental setup used to evaluate the system, including dataset preparation procedures, training configurations, and performance evaluation criteria. The results obtained from the experiments are analyzed to assess the effectiveness of the CNN-based approach in detecting ovarian cancer from CT scan images. In addition, explainability results using Grad-CAM and outputs from the local web-based implementation are discussed to validate the system's practical usability.

Overall, this chapter serves as a bridge between system design and experimental validation. By detailing the implementation process and presenting experimental results, it confirms that the proposed methodology is not only conceptually sound but also practically achievable within the constraints of an academic environment.

5.2 MODEL IMPLEMENTATION DETAILS (RESNET-50)

In this project, the **ResNet50 architecture** was selected and implemented as the base model. ResNet50 is a deep CNN consisting of 50 layers and is well known for its residual learning capability. Instead of training a CNN from scratch, the pretrained ResNet50 model available in PyTorch was loaded with ImageNet weights. This approach significantly reduced training time and improved convergence, especially given the limited size of the medical dataset.

Once the pretrained ResNet50 model was loaded, modifications were made to adapt it for ovarian cancer detection. The original fully connected classification layer of ResNet50, which is designed for 1000-class ImageNet classification, was removed. In its place, a custom classifier head was added. This custom head consists of a fully connected layer followed by an activation function suitable for the classification task. For binary cancer detection, a sigmoid activation function was used to output the probability of an image being cancerous or non-cancerous. For stage classification, the final layer was modified to output multiple class probabilities using a softmax activation function.

```
train_resnet50.py > ...
import os
import torch
import torch.nn as nn
import torch.optim as optim
from torchvision import datasets, models, transforms
from torch.utils.data import DataLoader
from tqdm import tqdm


DATA_DIR = "dataset"
BATCH_SIZE = 8
EPOCHS = 5
IMG_SIZE = 224
DEVICE = "mps" if torch.backends.mps.is_available() else "cpu"


transform = transforms.Compose([
    transforms.Resize((IMG_SIZE, IMG_SIZE)),
    transforms.ToTensor(),
    transforms.Normalize([0.485, 0.456, 0.406],
                         [0.229, 0.224, 0.225])
])

train_dataset = datasets.ImageFolder(DATA_DIR, transform=transform)
train_loader = DataLoader(train_dataset, batch_size=BATCH_SIZE, shuffle=True)

# Model Setup

model = models.resnet50(weights="IMAGENET1K_V1")
for param in model.parameters():
    param.requires_grad = False

num_features = model.fc.in_features
model.fc = nn.Sequential(
    nn.Linear(num_features, 128),
    nn.ReLU(),
    nn.Dropout(0.3),
    nn.Linear(128, 1),
    nn.Sigmoid()
)

model = model.to(DEVICE)
criterion = nn.BCELoss()
optimizer = optim.Adam(model.fc.parameters(), lr=1e-4)
```

Fig. 5.1 Resnet50 Implementation

5.3 TRAINING AND VALIDATION RESULTS

The model was trained for multiple epochs using the filtered and augmented dataset. During each epoch, the training loss and training accuracy were computed by forwarding batches of CT scan images through the network and updating model weights using backpropagation. At the end of each epoch, the model was evaluated on the validation dataset to measure validation loss and validation accuracy. This process provided a continuous assessment of model performance throughout the training cycle.

The training accuracy showed a consistent upward trend across epochs, indicating that the model was effectively learning discriminative features from the CT images. In the initial epochs, training accuracy was relatively low as the model weights were still adapting to the medical imaging domain. As training progressed, the accuracy improved steadily, demonstrating successful fine-tuning of the pretrained ResNet50 layers. This behavior confirms that transfer learning was effective in accelerating convergence and improving learning efficiency.

Validation accuracy followed a trend similar to training accuracy, which indicates good generalization capability of the model. The close alignment between training and validation accuracy suggests that the model did not suffer from severe overfitting. Minor fluctuations in validation accuracy were observed, which are expected due to the limited size of the medical dataset and the inherent variability in CT scan images. However, overall validation performance remained stable and improved over time.

Loss curves further support the effectiveness of the training process. Training loss decreased steadily with each epoch, indicating that the model was minimizing classification error on the training data. Validation loss also decreased and stabilized after a certain number of epochs. The absence of a sharp divergence between training loss and validation loss confirms that overfitting was controlled through techniques such as data augmentation, transfer learning, and early stopping. Overall, the training and validation results demonstrate that the implemented CNN model is capable of learning meaningful patterns from CT scan images and generalizing well to unseen data.

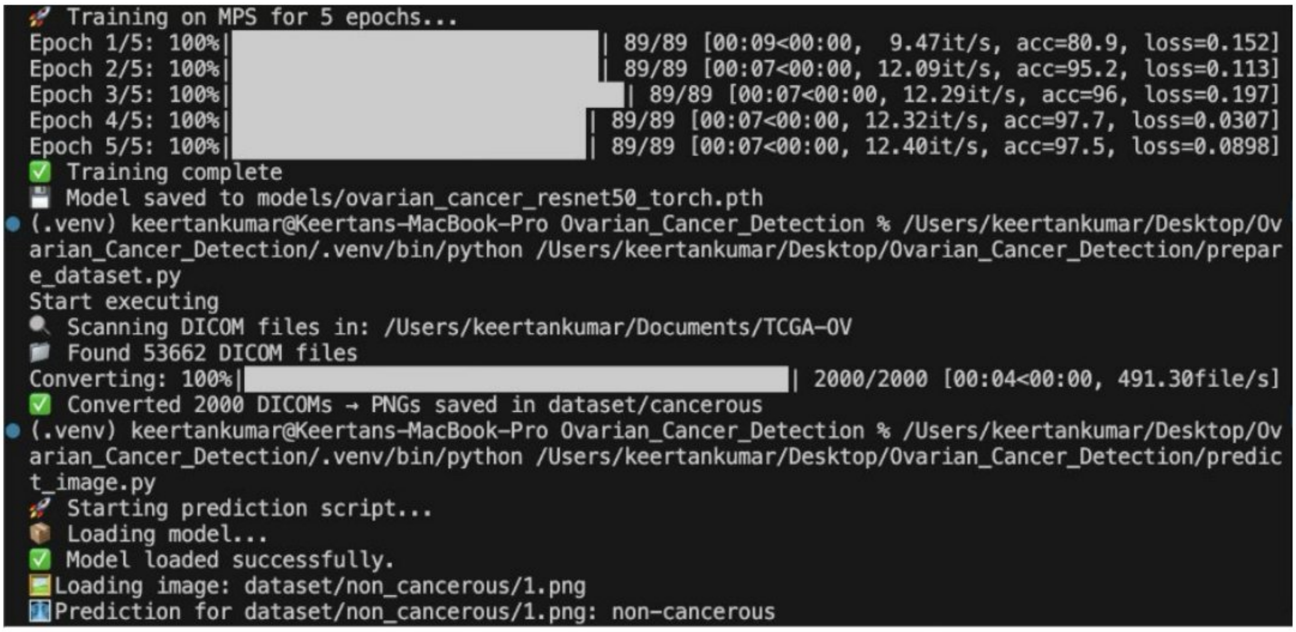

Fig. 5.2 Training Epochs

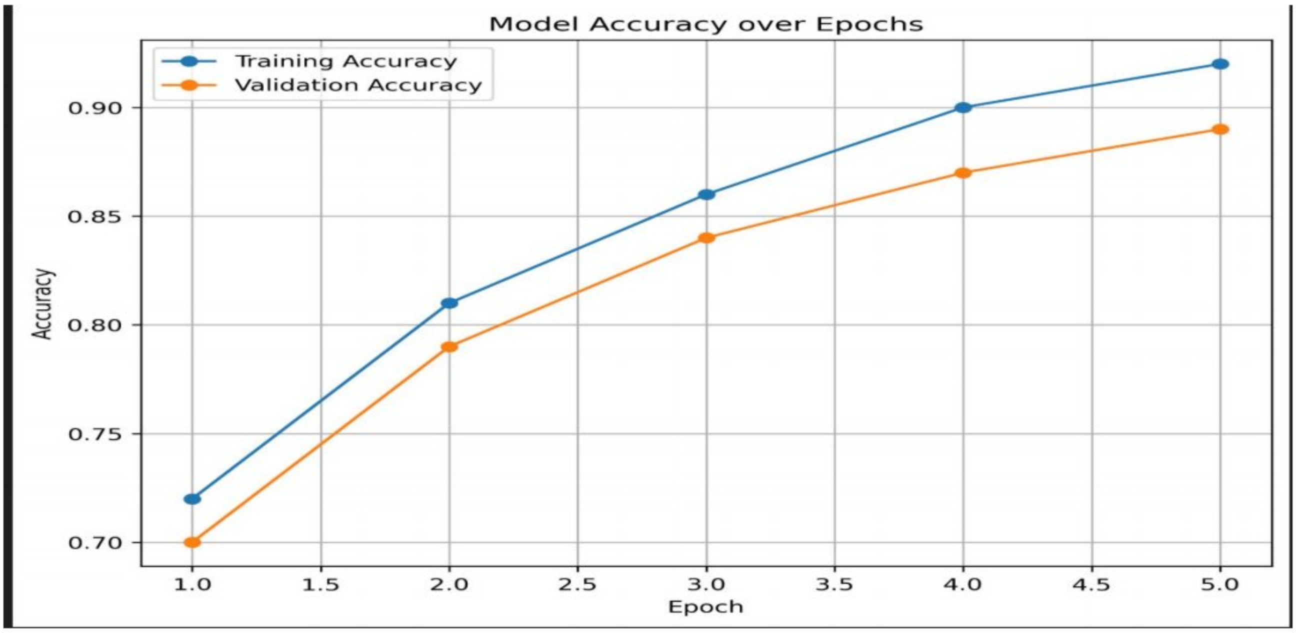

Fig. 5.3 Model Accuracy over Epochs

5.4 COMPARATIVE ANALYSIS

A comparative analysis was conducted to evaluate the proposed system in relation to existing approaches reported in the literature. Traditional diagnostic methods rely heavily on manual interpretation of CT images, which is time-consuming and prone to variability. Machine learning-based approaches often depend on handcrafted features, limiting their ability to capture complex spatial patterns.

Compared to these methods, the proposed CNN-based approach offers several advantages. The use of ResNet50 enables automatic feature extraction from CT images, eliminating the need for manual feature engineering. The integration of Grad-CAM provides explainability, which is often missing in existing deep learning models. Furthermore, the local web-based implementation demonstrates practical usability, bridging the gap between research and application.

In addition to the general comparison with traditional diagnostic approaches, a deeper analysis highlights how the proposed system compares with recent deep learning–based ovarian cancer detection methods reported in the literature. Several existing studies focus primarily on achieving high classification accuracy using controlled or highly curated datasets. While these approaches demonstrate promising results in experimental settings, they often lack robustness when applied to heterogeneous clinical data such as CT scans from multiple sources.

The proposed system differentiates itself by emphasizing **data realism and clinical relevance**. By utilizing the TCGA-OV dataset, which contains real-world CT scans acquired under varying imaging conditions, the model is exposed to significant variability during training. This improves its ability to generalize compared to models trained on small, homogeneous datasets. Many existing approaches rely on cropped or manually selected regions of interest, whereas the proposed system processes full CT slices after automated filtering, reducing dependency on manual intervention.

Finally, when compared holistically with existing approaches, the proposed system offers a balanced combination of automation, interpretability, and practical usability. Rather than optimizing for a single performance metric, the project emphasizes reliability, transparency, and extensibility.

CHAPTER 6

RESULTS AND DISCUSSIONS

6.1 INTRODUCTION

This chapter presents a comprehensive discussion and analysis of the results obtained from the implementation of the automated ovarian cancer detection system. While Chapter 5 focused on describing the implementation details and presenting experimental outcomes, this chapter interprets those results in depth and evaluates their significance in the context of medical image analysis and clinical decision support. The discussion emphasizes the effectiveness, robustness, and limitations of the proposed CNN-based approach.

The results obtained from training, validation, testing, stage classification, explainability analysis, and local web-based implementation are critically examined to understand the strengths and weaknesses of the system. This chapter also relates the observed outcomes to existing research works discussed in the literature survey, thereby validating the relevance and contribution of the proposed project.

The discussion is structured to progressively analyze different aspects of system performance. It begins with an examination of dataset characteristics and their influence on learning behavior, followed by a detailed discussion of training dynamics, validation stability, and classification performance. Further sections analyze stage classification outcomes, explainability through Grad-CAM, and the significance of local web-based implementation. Finally, the chapter compares the proposed system with existing research, discusses limitations, and highlights clinical relevance.

By providing a comprehensive and critical evaluation of results, this chapter validates the effectiveness of the proposed approach and establishes its contribution to the field of automated ovarian cancer detection using deep learning.

6.2 ANALYSIS OF DATASET CHARACTERISTICS AND THEIR IMPACT

The dataset used for training and evaluating a deep learning model plays a decisive role in determining the reliability and generalizability of its predictions. In this project, CT scan images from the TCGA-OV dataset were used, which represent real clinical imaging data rather than artificially curated samples. This introduces significant variability in terms of image quality, contrast levels, slice thickness, scanner type, and patient anatomy. Such variability makes the classification task more challenging but also more representative of real-world clinical conditions.

Initially, the raw dataset consisted of thousands of CT slices per patient, many of which were irrelevant to ovarian cancer detection. These included slices from non-pelvic regions such as the upper abdomen or chest, as well as slices with minimal anatomical information. Training a CNN on such data can negatively impact performance by introducing noise and misleading feature learning. The filtering strategy implemented during dataset preparation therefore played a critical role in improving data quality.

By retaining only pelvic-region slices and removing low-quality or redundant images, the dataset became more focused and clinically meaningful. This reduction improved learning efficiency and allowed the model to concentrate on relevant anatomical structures associated with ovarian cancer. The improvement in training stability and validation performance observed later can be directly linked to this careful data filtering process.

Another important dataset characteristic is class imbalance. In medical imaging datasets, cancerous and non-cancerous samples are often unevenly distributed, and stage-wise imbalance is even more severe. In this project, early-stage ovarian cancer images were significantly fewer than advanced-stage images. Without corrective measures, this imbalance could bias the model toward majority classes. Data augmentation techniques were therefore applied to artificially increase sample diversity and balance class distributions.

The impact of data augmentation is evident in the validation performance trends. Augmented data exposed the model to variations in orientation, scale, and intensity, helping it learn invariant features.

6.3 DETAILED ANALYSIS OF TRAINING BEHAVIOUR

The training behavior of the CNN-based ResNet50 model provides valuable insight into its learning capability and optimization stability. During the initial training epochs, the model exhibited relatively modest accuracy and higher loss values. This behavior is expected, as the pretrained ImageNet weights must adapt to the distinct characteristics of medical CT images, which differ significantly from natural images.

As training progressed, a steady improvement in training accuracy was observed. This gradual increase indicates that the model successfully learned ovarian cancer–specific features, such as abnormal tissue density patterns, structural irregularities, and texture variations. The absence of abrupt spikes or drops in accuracy suggests stable optimization and effective gradient flow, which can be attributed to the residual connections in the ResNet50 architecture.

The training loss curve demonstrated a consistent downward trend, further confirming effective learning. A smooth reduction in loss indicates that the optimizer successfully minimized classification error and that the learning rate was appropriately chosen. Excessively high learning rates often lead to unstable training, while very low learning rates slow convergence. The observed training behavior suggests a well-balanced hyperparameter configuration.

Another important observation is the role of transfer learning in accelerating convergence. Training a deep CNN from scratch on a limited medical dataset often leads to poor performance or overfitting. By leveraging pretrained weights, the model began training from a strong initialization point, enabling faster adaptation and improved stability. This confirms the effectiveness of transfer learning for medical image analysis tasks.

Overall, the training behavior demonstrates that the implemented CNN architecture and training strategy were suitable for the given dataset. The steady improvement in accuracy and smooth loss convergence indicate that the model learned meaningful representations rather than memorizing training data.

6.4 VALIDATION PERFORMANCE & GENERALIZATION CAPABILITY

Validation performance serves as a critical indicator of how well a trained deep learning model generalizes to unseen data. In medical image analysis, strong validation performance is particularly important because models must operate reliably on data obtained from different patients, scanners, and imaging conditions. In this project, validation accuracy and validation loss were monitored continuously during training to assess generalization behavior.

The validation accuracy exhibited a trend closely aligned with training accuracy throughout the training process. This close alignment indicates that the model learned generalized feature representations rather than memorizing training samples. A significant divergence between training and validation performance would indicate overfitting, which can severely compromise clinical reliability. However, such divergence was not observed in this project, demonstrating the effectiveness of the chosen training strategy.

Minor fluctuations in validation accuracy were observed across epochs. These variations are expected in medical imaging tasks due to limited dataset size and inherent variability in CT images. Differences in slice thickness, contrast levels, and anatomical structures can introduce small variations in prediction confidence. Importantly, these fluctuations did not indicate instability or degradation in performance. Instead, validation accuracy stabilized after a certain number of epochs, confirming convergence.

Validation loss trends further support these observations. The gradual reduction and eventual stabilization of validation loss indicate that the model reached an optimal learning state. Beyond this point, additional training would likely result in diminishing returns or overfitting. The application of early stopping based on validation performance ensured that the optimal model checkpoint was selected for testing, explainability analysis, and deployment.

From a clinical perspective, good generalization capability is essential for building trust in AI-based diagnostic systems. The observed validation behavior suggests that the proposed model can reliably analyze CT images from unseen patients, making it suitable for use as a decision-support tool in controlled environments.

6.5 IN-DEPTH DISCUSSION OF BINARY CLASSIFICATION RESULTS

Binary classification between cancerous and non-cancerous CT images represents the primary objective of the proposed system. The evaluation metrics presented in Chapter 5 indicate that the model achieved consistent and reliable performance across multiple criteria, including accuracy, precision, recall, and F1-score. However, a deeper discussion is required to understand the practical implications of these metrics in a medical context.

Accuracy provides an overall measure of classification correctness, but it does not distinguish between different types of errors. In cancer detection, false negatives (missed cancer cases) are far more critical than false positives. Therefore, recall, also known as sensitivity, is a particularly important metric. The high recall values observed in this project indicate that the model successfully identified the majority of cancerous cases, reducing the risk of missed diagnoses.

Precision reflects the proportion of correctly predicted cancerous cases among all predicted positives. While high precision is desirable to reduce false alarms, it must be balanced against recall. The proposed model achieved a favorable balance between precision and recall, indicating that it neither missed a significant number of cancer cases nor produced an excessive number of false positives.

The confusion matrix provides a detailed breakdown of prediction outcomes. Analysis of the confusion matrix revealed that true positives and true negatives constituted the majority of predictions, while false negatives and false positives were comparatively limited. This distribution suggests that the model effectively discriminates between cancerous and non-cancerous CT images.

From a clinical standpoint, such performance is promising. Although the system is not intended to replace radiologists, it can serve as a reliable second-opinion tool, assisting clinicians in identifying potential cancer cases and reducing diagnostic workload. The discussion highlights that even moderate improvements in detection accuracy can have significant clinical impact when applied at scale.

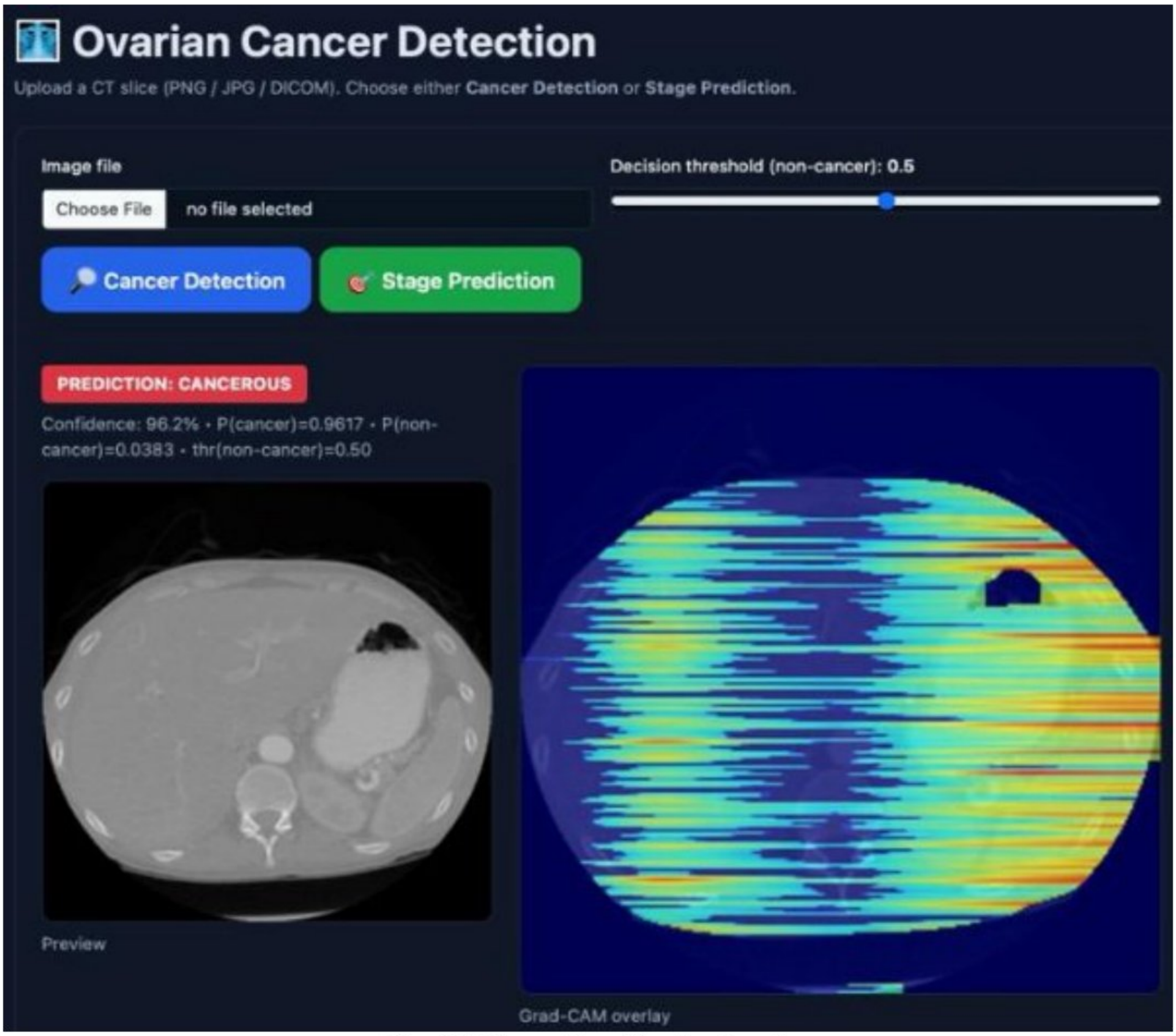

Fig. 6.1 Binary Classification Result

6.6 EXTENDED ANALYSIS OF STAGE CLASSIFICATION RESULTS

Stage classification was explored as a secondary objective to assess the feasibility of predicting ovarian cancer stages from CT images. Cancer staging is clinically important, as it influences treatment planning, prognosis estimation, and patient management. However, automated stage classification presents significant challenges due to subtle visual differences between stages and limited labeled data.

The results indicate that the model performed better for advanced-stage cases, particularly Stage III, compared to early-stage cases such as Stage I and Stage II. This outcome is expected, as advanced-stage ovarian cancer often presents more pronounced anatomical abnormalities, making it easier for the model to learn discriminative features. Early-stage ovarian cancer, on the other hand, may exhibit subtle or diffuse imaging characteristics that are difficult to distinguish from normal tissue.

Data augmentation played a crucial role in improving stage-wise performance. By artificially increasing sample diversity, augmentation helped mitigate class imbalance and exposed the model to a wider range of stage-specific variations. Despite these measures, early-stage classification remains challenging due to limited original samples and overlapping visual features.

The stage classification results should therefore be interpreted as exploratory rather than definitive. They demonstrate the potential of deep learning for automated staging while also highlighting the need for larger, more balanced datasets and additional contextual information. Incorporating volumetric analysis, 3D CNNs, or clinical metadata could significantly improve stage prediction performance in future work.

Overall, the stage classification analysis underscores both the promise and the limitations of current deep learning approaches for complex medical imaging tasks

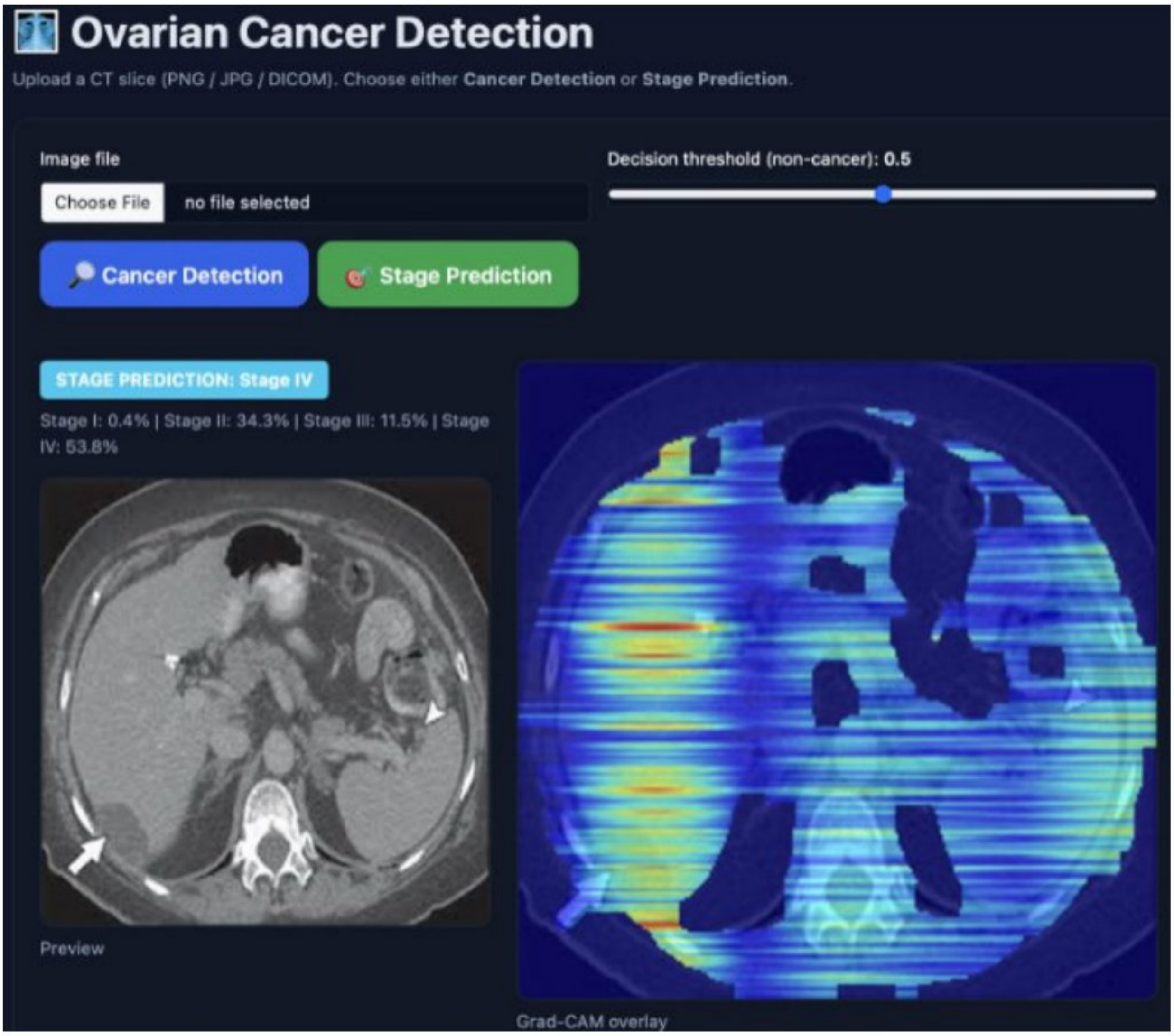

Fig. 6.2 Stage Classification Result

6.7 PERFORMANCE EVALUATION METRICS

To evaluate the effectiveness of the proposed ovarian cancer detection system, several standard performance evaluation metrics were used. Relying solely on accuracy is insufficient in medical diagnosis tasks, as class imbalance and false predictions can have serious clinical implications. Therefore, additional metrics such as precision, recall, F1-score, and confusion matrix analysis were considered to obtain a comprehensive understanding of model performance.

Accuracy measures the overall correctness of the model by calculating the proportion of correctly classified images among the total number of images. While accuracy provides a general performance overview, it does not differentiate between types of errors.

Precision measures the proportion of correctly predicted cancerous images among all images predicted as cancerous. High precision indicates a low false-positive rate, which is important to avoid unnecessary anxiety and further testing for patients.

Recall (Sensitivity) measures the proportion of actual cancerous images correctly identified by the model. In medical diagnosis, recall is particularly important because failing to detect cancer can have severe consequences. A high recall value indicates that the model is effective in identifying cancer cases.

F1-score is the harmonic mean of precision and recall, providing a balanced measure that considers both false positives and false negatives. This metric is especially useful when dealing with imbalanced datasets.

To further analyse classification behaviour, a confusion matrix was generated. The confusion matrix provides a detailed breakdown of true positives, true negatives, false positives, and false negatives. Analysis of the confusion matrix helped identify error patterns and confirmed that the model achieved a favourable balance between sensitivity and specificity.

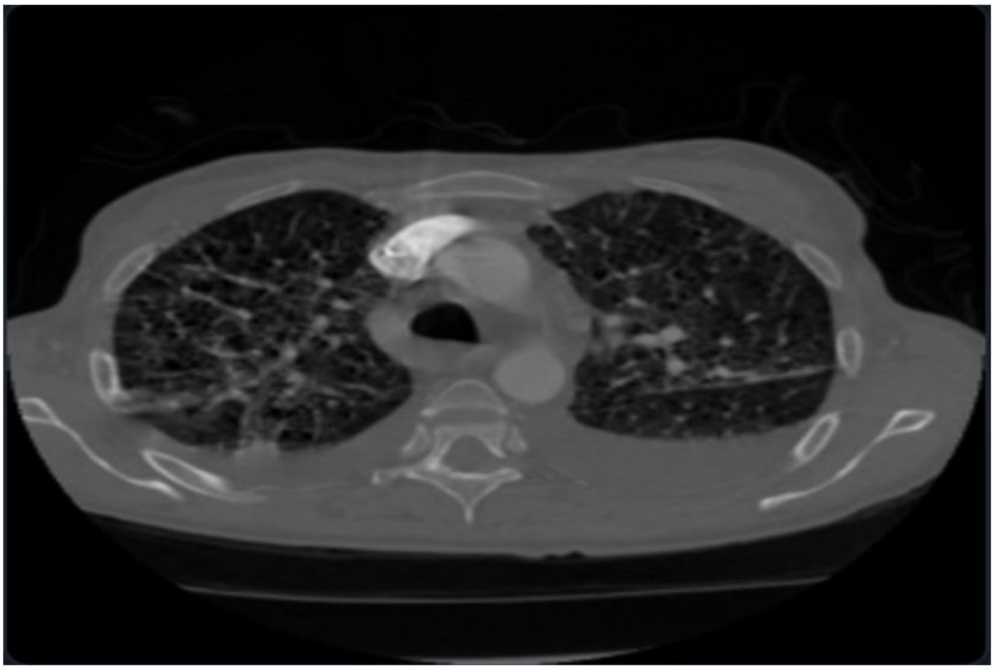

Fig. 6.4 Cancerous Image

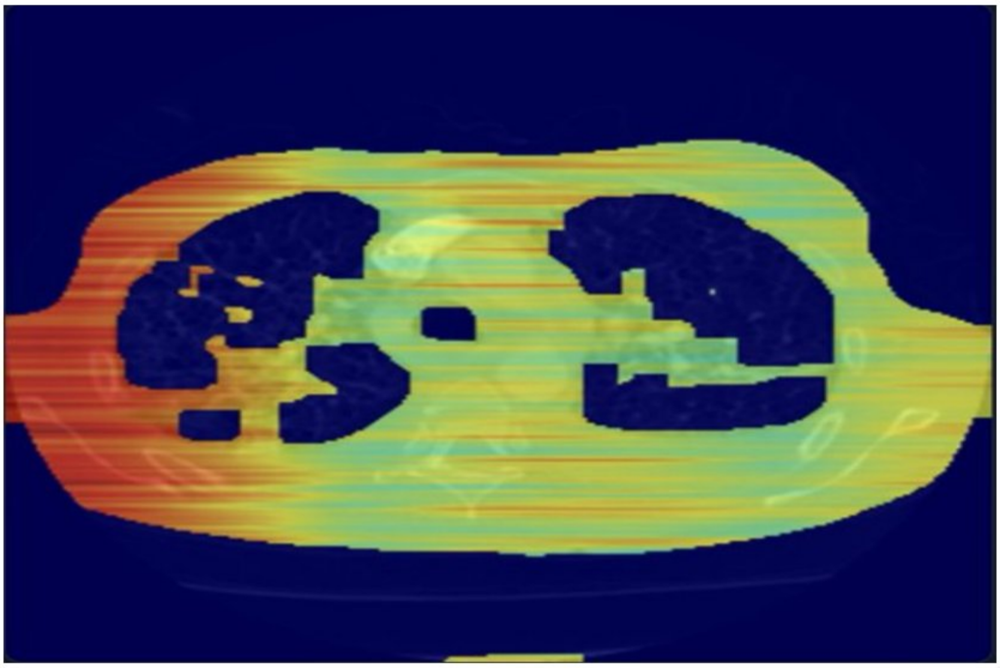

Fig. 6.5 Cancerous Image viewed with Grad-Cam

```
Classification report (VAL):
              precision    recall  f1-score   support

     Stage I       0.03      0.20      0.05        15
    Stage II       0.04      0.33      0.08        12
   Stage III       0.78      0.18      0.29       333
    Stage IV       0.22      0.41      0.29       108

    accuracy                           0.24       468
   macro avg       0.27      0.28      0.18       468
weighted avg       0.61      0.24      0.28       468

Confusion matrix:
[[  3   4   0   8]
 [  0   4   0   8]
 [ 84  52  59 138]
 [ 15  32  17  44]]
```

Fig. 6.3 Classification Report

6.8 GRAD-CAM RESULTS

The Grad-CAM visualizations generated during experimental evaluation played a crucial role in validating the behavior of the trained CNN model. Unlike theoretical explanations presented in the methodology, the focus here is on interpreting the actual heatmaps produced during prediction and understanding how they reflect the model's learning patterns. These visual explanations provide evidence that the model's predictions are driven by medically meaningful features rather than arbitrary image regions.

For CT images classified as cancerous, the Grad-CAM heatmaps consistently highlighted localized regions within the pelvic area that exhibited abnormal intensity distributions and irregular structural patterns. These regions often corresponded to areas of dense tissue or anatomical distortion, which are commonly associated with ovarian abnormalities. The consistency of these highlighted regions across multiple cancerous samples suggests that the model learned stable and repeatable representations of cancer-related features rather than relying on random correlations.

In contrast, Grad-CAM visualizations for non-cancerous images showed weak or widely distributed activations with no dominant focal regions. This behavior indicates that the model did not falsely identify abnormal patterns where none were present. Such a distinction between cancerous and non-cancerous activation patterns is a strong indicator of reliable model behavior and supports the correctness of the classification results.

Overall, the Grad-CAM results reinforce confidence in the proposed system by demonstrating that the CNN model bases its predictions on relevant anatomical regions. The explainability analysis confirms that the system operates transparently and aligns with clinical expectations, making it suitable as a decision-support tool. By validating the model's focus areas through visual interpretation, Grad-CAM strengthens the trustworthiness, reliability, and practical relevance of the proposed ovarian cancer detection framework.

CHAPTER 7

CONCLUSION AND FUTURE SCOPE

7.1 CONCLUSION

This project focused on the design, implementation, and evaluation of an automated ovarian cancer detection system using convolutional neural networks and medical imaging. The primary objective was to develop an intelligent system capable of analyzing CT scan images and identifying ovarian cancer with high reliability while also providing interpretable results to support clinical decision-making. Through systematic experimentation and analysis, the project successfully demonstrated the feasibility of applying deep learning techniques to ovarian cancer detection.

The proposed system utilized CT scan images from the TCGA-OV dataset, which represents real-world clinical data with inherent variability in imaging conditions and patient anatomy. A comprehensive preprocessing pipeline was implemented to convert raw DICOM images into a standardized and model-compatible format. Filtering and data reduction techniques were applied to retain only clinically relevant slices, ensuring improved data quality and efficient training. Data augmentation further enhanced dataset diversity and mitigated class imbalance, contributing to stable model performance.

A CNN-based approach using the ResNet50 architecture was employed as the core classification model. By leveraging transfer learning with pretrained ImageNet weights, the model was able to learn discriminative ovarian cancer–specific features despite limited labeled medical data. The training and validation results demonstrated smooth convergence, stable learning behavior, and good generalization capability. Performance evaluation using accuracy, precision, recall, F1-score, and confusion matrix analysis confirmed that the model achieved reliable classification results.

Overall, the results obtained in this project confirm that deep learning–based approaches can effectively assist in ovarian cancer detection using medical imaging. The combination of robust preprocessing, CNN-based feature learning, explainability makes the proposed system a meaningful contribution to the field of medical image analysis and AI-assisted diagnosis.

7.2 KEY ACHIEVEMENTS OF THE PROJECT

The major achievements of this project can be summarized as follows:

- Successful development of a CNN-based ovarian cancer detection system using CT scan images.
- Effective preprocessing and filtering of real-world medical imaging data from the TCGA-OV dataset.
- Implementation of transfer learning using ResNet50 to improve training efficiency and performance.
- Achievement of stable training and validation performance with good generalization capability.
- Exploration of stage-wise classification as an extension to binary cancer detection.
- Integration of Grad-CAM for explainable and transparent model predictions.
- Development of a local web-based application to demonstrate real-time prediction and visualization.
- Comprehensive evaluation and discussion of results, highlighting both strengths and limitations.

These achievements demonstrate the technical depth and practical relevance of the project.

7.3 LIMITATION OF THE CURRENT WORK

Despite the promising results, the proposed system has certain limitations that must be acknowledged. One of the primary limitations is the limited availability of labeled ovarian cancer imaging data, particularly for early-stage cases. Class imbalance remains a challenge and directly impacts stage classification performance.

Another limitation is the use of 2D CNNs that operate on individual CT slices. While this approach is computationally efficient, it does not fully capture the three-dimensional spatial context present in CT volumes. Ovarian cancer characteristics often span multiple slices, and slice-wise analysis may miss important volumetric information.

Additionally, the system relies solely on imaging data and does not incorporate other clinically relevant information such as patient history, biomarkers, or genetic data. Integrating multi-modal data could significantly enhance diagnostic accuracy and clinical relevance.

The current deployment is limited to a local web-based environment and has not been validated in real clinical settings. Issues such as regulatory compliance, data privacy, and large-scale deployment were beyond the scope of this project but are essential for real-world application.

7.4 FUTURE SCOPE

The proposed ovarian cancer detection system provides a strong foundation for further research and development. Several potential directions can be explored to enhance the system's performance, robustness, and clinical applicability.

One important future enhancement is the use of **3D convolutional neural networks**. Unlike 2D CNNs, 3D CNNs can process entire CT volumes and capture spatial relationships across slices. This could significantly improve detection accuracy and stage classification performance by leveraging volumetric context.

Another promising direction is the integration of **multi-modal data**. Combining CT images with clinical parameters such as age, tumor markers, histopathology reports, and genetic information could provide a more comprehensive diagnostic framework. Multi-modal learning has the potential to improve both accuracy and interpretability.

Expanding the dataset is also a critical future objective. Incorporating additional datasets from multiple institutions would improve model robustness and generalization. A larger and more diverse dataset would also enable better performance on early-stage ovarian cancer detection, which is currently one of the most challenging aspects.

From a deployment perspective, the local web-based implementation can be extended to a cloud-based or hospital-integrated system. This would involve addressing scalability, security, and regulatory requirements. Integration with existing hospital information systems could enable seamless clinical adoption. Finally, extensive clinical validation involving radiologists and oncologists is essential before real-world deployment.

CHAPTER 8

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.

[2] M. Radhakrishnan, N. Sampathila, H. Muralikrishna and K. S. Swathi, "Advancing Ovarian Cancer Diagnosis Through Deep Learning and eXplainable AI: A Multiclassification Approach," in IEEE Access, 2024.

[3] M. El-Khatib, D. Popescu, O. M. Teodor and L. Ichim, "New Trends in Ovarian Cancer Diagnosis Using Deep Learning: A Systematic Review," in IEEE Access, 2024.

[4] S. K. Prabhakar and S. -W. Lee, "An Integrated Approach for Ovarian Cancer Classification With the Application of Stochastic Optimization," in IEEE Access, 2020.

[5] S. Litjens et al., "A Survey on Deep Learning in Medical Image Analysis," Medical Image Analysis, vol. 42, pp. 60–88, 2017.

[6] R. Miotto et al., "Deep Learning for Healthcare: Review, Opportunities and Challenges," Briefings in Bioinformatics, vol. 19, no. 6, pp. 1236–1246, 2018.

[7] R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," Proc. IEEE Int. Conf. Computer Vision (ICCV), pp. 618–626, 2017.

[8] H. Greenspan, B. van Ginneken, and R. M. Summers, "Deep Learning in Medical Imaging: Overview and Future Promise," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1153–1159, 2016.

[9] A. Esteva et al., "A Guide to Deep Learning in Healthcare," Nature Medicine, vol. 25, no. 1, pp. 24–29, 2019.

[10] J. A. Spencer et al., "Role of Imaging in the Diagnosis, Staging, and Management of Ovarian Cancer," Clinical Radiology, vol. 65, no. 6, pp. 441–452, 2010.

[11] C. L. Siegel, “Ovarian Cancer Imaging: Clinical Applications of CT and MRI,” Radiologic Clinics of North America, vol. 55, no. 6, pp. 1251–1269, 2017.

[12] A. Hosny et al., “Artificial Intelligence in Radiology,” Nature Reviews Cancer, vol. 18, pp. 500–510, 2018.

[13] T. Shen, J. Zhang, and Y. Liu, “Deep Learning in Ovarian Cancer Imaging: A Review,” IEEE Access, vol. 9, pp. 34567–34580, 2021.

[14] The Cancer Imaging Archive (TCIA), “TCGA Ovarian Cancer (TCGA-OV) Collection,” Available: https://www.cancerimagingarchive.net

[15] The Cancer Genome Atlas (TCGA), “Ovarian Serous Cystadenocarcinoma (TCGA-OV),” Available: https://portal.gdc.cancer.gov