

BLOCK - A MULTI CANCER DISEASE DETECTION
A PROJECT REPORT

Submitted by

**TALLURU VENKATA BHUVANESH [RA2011003011298]
NANDIMANDALAM VIVEK [RA2011003011302]**

Under the Guidance of
Dr . R. THENMOZHI

Associate Professor, Department of Computing Technologies

In partial fulfillment of the Requirements for the Degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE ENGINEERING



**DEPARTMENT OF COMPUTING TECHNOLOGIES
COLLEGE OF ENGINEERING AND TECHNOLOGY
SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR – 603203**

MAY 2024



SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR – 603 203

BONAFIDE CERTIFICATE

Certified that 18CSP109L project report titled "**BLOCK - A MULTI CANCER DISEASE DETECTION**" is the Bonafide work of "**TALLURU VENKATA BHUVANESH [RA2011003011298], NANDIMANDALAM VIVEK [RA2011003011302]**" who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion for this or any other candidate.

R. Thenmozhi
Dr. R. THENMOZHI
SUPERVISOR
Associate Professor
Department of Computing Technologies

Dr. A. Pandian
Dr. A. PANDIAN
PANEL HEAD
Associate Professor
Department of Computing Technologies

M. Pushpalatha

Dr. M. PUSHPALATHA
HEAD OF THE DEPARTMENT
Department of Computing Technologies



INTERNAL EXAMINER

EXTERNAL EXAMINER



Department of Computing Technologies
SRM Institute of Science & Technology

Own Work Declaration Form

Degree/ Course : Bachelor of Technology, Computer Science Engineering

Student Names : Talluru Venkata Bhuvanesh, Nandimandalam Vivek

Registration Numbers : RA2011003011298, RA2011003011302

Title of Work : BLOCK - A MULTI CANCER DISEASE DETECTION.

We hereby certify that this assessment complies with the University's Rules and Regulations relating to Academic misconduct and plagiarism, as listed in the University Website, Regulations, and the Education Committee guidelines.

We confirm that all the work contained in this assessment is my / our own except where indicated, and that we have met the following conditions:

- Clearly references / listed all sources as appropriate
- Referenced and put in inverted commas all quoted text (from books, web, etc)
- Given the sources of all pictures, data etc. that are not my own
- Not made any use of the report(s) or essay(s) of any other student(s) either past or present
- Acknowledged in appropriate places any help that I have received from others (e.g. fellow students, technicians, statisticians, external sources)
- Compiled with any other plagiarism criteria specified in the Course handbook / University website

I understand that any false claim for this work will be penalized in accordance with the University policies and regulations.

DECLARATION:

I am aware of and understand the University's policy on Academic misconduct and plagiarism and I certify that this assessment is my / our own work, except where indicated by referring, and that I have followed the good academic practices noted above.

Student 1 Signature:

Student 2 Signature:

Date:

If you are working in a group, please write your registration numbers and sign with the date for every student in your group.

ACKNOWLEDGEMENT

We express our humble gratitude to **Dr. C. Muthamizhchelvan**, Vice-Chancellor, SRM Institute of Science and Technology, for the facilities extended for the project work and his continued support.

We extend our sincere thanks to **Dr. T. V. Gopal**, Dean-CET, SRM Institute of Science and Technology, for his invaluable support.

We wish to thank **Dr. Revathi Venkataraman**, Professor and Chairperson, School of Computing, SRM Institute of Science and Technology, for her support throughout the project work.

We are incredibly grateful to our Head of the Department, **Dr. M. Pushpalatha**, Professor, Department of Computing Technologies, SRM Institute of Science and Technology, for her suggestions and encouragement at all the stages of the project work.

We want to convey our thanks to our program coordinators, **Dr. S. Godfrey Winster**, Associate Professor, **Dr. M. Baskar**, Associate Professor, **Dr. P. Murali**, Associate Professor, **Dr. J. Selvin Paul Peter**, Associate Professor, **Dr. C. Pretty Diana Cyril**, Assistant Professor and **Dr. G. Padmapriya**, Assistant Professor, Panel Head, **Dr. A. Pandian** Associate Professor and Panel Members, **Dr. K. Priyadarshini** Assistant Professor, **Dr. P. Velmurugan** Assistant Professor and **Dr. R. Thenmozhi** Associate Professor Department of Computing Technologies, SRM Institute of Science and Technology, for their inputs during the project reviews and support.

We register our immeasurable thanks to our Faculty Advisor, **Dr. N. Nithyanandam**, Assistant Professor, Department of Computing Technologies, SRM Institute of Science and Technology, for leading and helping us to complete our course.

Our inexpressible respect and thanks to our guide, , **Dr. R. Thenmozhi**, Associate Professor Department of Computing Technologies, SRM Institute of Science and Technology, for providing me with an opportunity to pursue my project under her mentor-ship. She provided me with the freedom and support to explore the research topics of my interest. Her passion for solving problems and making a difference in the world has always been inspiring.

We sincerely thank all the staff and students of Computing Technologies Department, SRM Institute of Science and Technology, for their help during our project. Finally, we would like to thank parents, family members, and friends for their unconditional love, constant support, and encouragement.

Talluru Venkata Bhuvanesh[RA2011003011298]

Nandimandalam Vivek[RA2011003011302]

ABSTRACT

This abstract provides an overview of the application of deep learning techniques for the detection of cancer diseases: cervical cancer, lung and colon cancer, oral cancer, kidney cancer, breast cancer, and brain cancer. Deep learning has gained significant attention in recent years due to its ability to automatically learn and extract complex patterns from large datasets, enabling accurate and efficient cancer detection. For cervical cancer detection, deep learning models have been developed to analyse Pap smear images and cervical cell morphology, aiding in early diagnosis and treatment. Similarly, in lung and colon cancer detection, deep learning algorithms have demonstrated promising results in the analysis of medical images, such as CT scans and colonoscopy images, assisting in the identification of cancerous regions and providing prognostic information. In the case of oral cancer, deep learning techniques have been used to analyse histopathological images, offering accurate classification and segmentation of oral tissues, allowing for early detection and intervention. For kidney cancer, deep learning models have been employed to analyse medical imaging data, such as MRI and ultrasound, facilitating the detection and characterization of tumors. Breast cancer detection has also seen significant advancements with deep learning, where models have been developed to analyse mammography and histopathology images, improving the accuracy of diagnosis and reducing false negatives. Lastly, deep learning approaches have been explored for brain cancer detection using MRI and PET imaging data, contributing to the detection and segmentation of tumor regions, assisting in treatment planning and monitoring. In summary, the application of deep learning techniques for the detection of cervical, lung and colon, oral, kidney, breast, and brain cancers has shown promising results in various imaging modalities, enabling more accurate diagnosis and timely intervention in the battle against these diseases.

TABLE OF CONTENTS

ABSTRACT	vi
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	x
1 INTRODUCTION	1
1.1 Overview	1
1.2 Objectives	2
1.3 Problem statement	3
2 LITERATURE SURVEY	5
3 SYSTEM ARCHITECTURE AND DESIGN	11
3.1 General	11
3.2 System Architecture Diagram	11
3.3 Development Environment	13
3.3.1 Hardware Requirements	13
3.3.2 Software Requirements	14
3.4 Design of the Entire System	14
3.4.1 Data Flow Diagram	14
3.4.2 Use Case Diagram	17
3.4.3 Sequence Diagram	18
3.4.4 Activity Diagram	19
3.5 Existing System	20
3.5.1 Drawbacks in Existing System	21
4 METHODOLOGY	24
4.1 Modules	25
4.1.1 Imaging Modules	25
4.1.2 Molecular Diagnostics Module	26
4.1.3 Artificial Intelligence Analytics Module	27
5 RESULTS AND DISCUSSION	29
6 CONCLUSION AND FUTURE ENHANCEMENT	36
7 REFERENCES	42
APPENDIX 1	44
APPENDIX 2	69
PLAGIARISM REPORT	71
PAPER PUBLICATION PROOF	80

LIST OF TABLES

3.3.1	Hardware Requirements.....	13
3.3.2	Software Requirements.....	14

LIST OF FIGURES

3.2	Architecture Diagram.....	11
3.4.1	Data Flow Diagram	15
3.4.2	Use case Diagram	17
3.4.3	Sequence Diagram	18
3.4.4	Activity Diagram	19
5.1	Accuracy rates for Cervical cancer	29
5.2	Accuracy rates for brain cancer	30
5.3	Accuracy rates for Breast cancer	30
5.4	Accuracy rates for Kidney cancer	31
5.5	Accuracy rates for lung and colon cancer	32
5.6	Accuracy rates for Lymphoma	32
5.7	Accuracy rates for Oral cancer	33
5.8	Accuracy rates for ML Algorithms	34

LIST OF ABBREVIATIONS

CBCIS	Cancer Body Check Identification System
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
CT	Computed Tomography
DNASeq	DNA Sequencing
AI-ML	Artificial Intelligence and Machine Learning
NGS	Next-Generation Sequencing
EMR	Electronic Medical Records
PCR	Polymerase Chain Reaction
IHC	Immunohistochemistry

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

One of the main causes of mortality worldwide is cancer, and better patient outcomes and effective treatment depend on early identification. A subset of machine learning called deep learning has shown promise in the identification and treatment of certain cancers. The six forms of cancer that will be covered in this paragraph are kidney, breast, brain, lung, and colon cancers, as well as oral and cervical cancers, and how deep learning algorithms may help identify them.

Cervical cancer primarily affects women and is caused by the human papillomavirus (HPV). Deep learning models can analyse cervical cells and detect abnormalities that may indicate the presence of cancer. By training models on large dataset of cervical images, algorithms can accurately identify cancerous cells and help in early detection.

Lung and colon cancer are prevalent and deadly forms of cancer. Deep learning algorithms are capable of identifying possible malignancies and lesions in the lungs and intestines by analyzing medical pictures such as CT scans and X-rays. By identifying patterns and anomalies in these images, these algorithms can aid radiologists in detecting cancer at earlier stages and in making more accurate diagnoses.

Oral cancer refers to cancer that affects the mouth and throat. Deep learning models can be trained on images of the oral cavity and utilize pattern recognition techniques to identify tumors and lesions in this area. By assisting dentists and oral surgeons in detecting early signs of cancer, these algorithms can potentially save lives and improve patient outcomes.

Kidney cancer is often difficult to diagnose at early stages due to the lack of specific symptoms. In order to detect possible kidney cancers, deep learning algorithms can assist in the analysis of medical imaging such as CT and MRI images. By flagging suspicious areas, these algorithms can assist radiologists in making more accurate diagnoses and determining appropriate treatment plans.

Globally, breast cancer is the most frequent cancer to affect women. Deep learning algorithms can analyze mammograms and identify potential tumor or areas of concern. By

utilizing advanced image processing techniques, these algorithms can aid radiologists in detecting breast cancer at earlier stages, leading to better treatment options and improved survival rates.

Brain cancer refers to cancers that originate in the brain or its surrounding tissues. Deep learning algorithms can analyze medical imaging such as MRI scans to detect tumors or abnormalities in the brain. By accurately identifying these markers, these algorithms can help physicians develop tailored treatment plans and potentially improve patient outcomes.

To sum up, deep learning algorithms hold the potential to transform the process of detecting and diagnosing cancer, regardless of the type of cancer. These algorithms use artificial intelligence to help medical personnel detect malignant cells and tumors early on, which could result in more successful therapies and better patient outcomes.

1.2 OBJECTIVES

The objectives of implementing an All Body Cancer Check Identification System (ABCIS) are multifaceted and aimed at addressing the challenges inherent in current cancer detection and diagnosis processes while working to enhance patient outcomes, expedite the provision of healthcare, and promote medical research. The following lists the main goals of ABCIS:

Early Detection: One primary objective of ABCIS is to enable the early detection of cancer across multiple organs and tissues. Early detection is critical for improving patient prognosis and increasing the success rates of treatment interventions. By integrating various screening modalities, including imaging scans and molecular tests, ABCIS aims to identify cancerous growths at their earliest stages when they are most treatable.

Comprehensive Screening: ABCIS seeks to provide a comprehensive screening solution that encompasses the detection of cancer affecting the human body. By offering a unified approach to cancer screening, ABCIS aims to minimize the risk of missed diagnoses and ensure that patients receive thorough evaluations for various malignancies, thereby enhancing overall healthcare quality and patient safety.

Accuracy and Precision: ABCIS is designed to prioritize accuracy and precision in cancer detection and diagnosis. Leveraging advanced imaging technologies, molecular assays, and artificial intelligence algorithms, the system aims to minimize false positives and false negatives, thereby reducing unnecessary interventions and ensuring that patients receive timely and appropriate care based on accurate diagnostic information.

Personalized Medicine: Another objective of ABCIS is to facilitate personalized medicine approaches tailored to individual patient profiles. By integrating patient-specific data, such as genetic information, medical history, and imaging findings, ABCIS enables healthcare providers to develop personalized treatment plans that optimize therapeutic outcomes while minimizing adverse effects and treatment-related complications.

Streamlined Workflow: ABCIS aims to streamline the cancer diagnosis workflow by consolidating multiple screening modalities and diagnostic tests into a cohesive and integrated platform. By automating data analysis, report generation, and communication between healthcare providers, ABCIS reduces administrative burdens and enhances operational efficiency, allowing clinicians to focus more on patient care and decision-making.

Interoperability and Data Sharing: Ensuring interoperability and seamless data sharing between healthcare systems and providers is a key objective of ABCIS. By integrating with electronic medical record systems and health information exchanges, ABCIS enables the secure exchange of patient information, imaging results, and diagnostic reports, facilitating collaborative decision-making and continuity of care across different healthcare settings.

Research and Population Health: ABCIS aims to support research initiatives aimed at advancing our understanding of cancer biology, risk factors, and treatment outcomes. By aggregating de-identified patient data and imaging studies, ABCIS enables population-level analytics and epidemiological research, helping to identify trends, disparities, and areas for targeted intervention to improve population health outcomes.

In conclusion, the objectives of implementing an All Body Cancer Check Identification System are rooted in improving cancer detection, diagnosis, and treatment outcomes through early detection, comprehensive screening, accuracy, precision, personalized medicine, streamlined workflow, interoperability, and data sharing. By addressing these objectives, ABCIS has the potential to transform the landscape of cancer care, enhance patient outcomes, and contribute to advancements in medical research and population health.

1.2 PROBLEM STATEMENT

The development of this project aims to revolutionize cancer detection and diagnosis by offering a comprehensive solution for identifying six types of cancer that commonly occur in the human body. Cancer, a complex and multifaceted disease, presents significant challenges in its early detection and accurate diagnosis, often leading to delayed treatment and poorer outcomes

for patients. The need for a unified and efficient approach to cancer screening across various organs and tissues has become increasingly evident in modern healthcare.

Existing cancer detection methods often focus on specific types of cancer or particular regions of the body, resulting in fragmented diagnostic processes that may overlook the presence of other malignancies.

This project aims to address these challenges by offering a holistic and integrated solution for cancer screening and identification. By leveraging advancements in medical imaging technology, molecular biology, and artificial intelligence, the system provides a comprehensive assessment of a patient's cancer risk across multiple organs and tissues. Through a combination of cutting-edge molecular methods like DNA sequencing and immunohistochemistry in conjunction with non-invasive imaging modalities like PET, CT, and MRI, the ABCIS provides highly accurate early cancer detection and characterization of a variety of cancer types.

One of the key features of this project is its incorporation of AI and machine learning algorithms to analyse complex medical data and identify patterns indicative of cancerous growths. By training these algorithms on large datasets of imaging scans, genetic profiles, and clinical outcomes, the system can learn to recognize subtle signs of malignancy that may be missed by human observers. This not only enhances the sensitivity and specificity of cancer detection but also streamlines the diagnostic process by providing timely and actionable insights to healthcare providers.

Furthermore, this project integrates with electronic medical record systems to facilitate seamless data exchange and collaboration among healthcare professionals. By consolidating patient information, imaging results, and diagnostic reports in a centralized platform, the system ensures continuity of care and enables personalized treatment planning based on individual patient profiles.

This interoperability also enables population-level analytics and research, allowing healthcare organizations to identify trends, risk factors, and treatment outcomes across diverse patient populations.

In conclusion, the development of this project represents a significant advancement in cancer screening and diagnosis, offering a comprehensive solution for detecting eight types of cancer that commonly affect the human body.

CHAPTER 2

LITERATURE SURVEY

N. V. Orlov, et al., “Automatic classification of lymphoma ideas following reconstruct situated general face,” IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 4, pp. 1003–1013, Jul. 2010. This study looks on the machinelike classification of cloak can lymphoma, follicular lymphoma, and never-ending lymphocytic leukemia—three frequent forms of unhealthy lymphoma. The aim having to do with this study follow identify patterns that would outweigh the type-imaginative categorization of lymphoma malignancies, apart from the definable writing of picture content through the use of manipulative dream plans. In this study, a novel two-stage method was used. Raw pixels were converted into spooky planes at the extrinsic level resorting to one renewals. The calculation difficult determining the organic (Fourier, Chebyshev, and wavelets) and complex (Chebyshev of Fourier and wavelets of Fourier) revamps. The second stage (the principal level) captured nakedness pixels and ghostly planes therefore forward [1].

Cho, J., et al., “Using CT/MRI ideas and information outside disregarding excite deep education models, various types of virulence possibly highly classified,” IEEE Access, vol. 11, pp. 10336–10354, 2023. This study delves into the field of deep instruction for malignancy classification resorting to CT and MRI ideas. It tackles a critical challenge in this place place rule, that is to say the stop of model disregarding. When employing deep instruction models for cyst classification, it's possessed by guarantee that they forbiddance drop former calm news concurrently with an activity the development process. The paper debates methods and approaches to insist the model's truth by expediting this issue. By accomplishment so, it donates to the occurrence of more trustworthy and forceful deep instruction models for the classification of various types of virulence settled restorative image[2].

Gupta, S., et al., “Prediction Performance of Deep Learning for Colon Cancer Survival Prediction on SEER Data,” BioMed Research International, vol. 2022, Article ID 1467070, pp. 1–12, 2022. This research investigates the calling effectiveness of deep information when used to the task of foreseeing maintenance rates for things acknowledged following colon lump. Leveraging SEER file, the study aims to decorate the veracity of lastingness forecasts. Accurate addition forecasts play a fault-finding act in conceiving custom-made position plans and reconstructing

patient effects. By determining the conduct of deep information models in this place place footing, the study donates valuable judgments that protect aid clinicians in making more experienced decisions and ultimately cause success better patient protect those management colon malignancy [3].

Alanazi, S. A., and so et al., "Boosting breast virulence finding applying convolutional pertain whole," J. Healthcare Eng., vol. 2021, pp. 1-11, Apr. 2021. This study focuses on conscience Cancer, that is to say understanding usual a fairly frequent type of resentment in daughters and springs in shame cartons. After body part sac, feelings cyst poses a difficult risk to a she's tumor. This study plans allocating convolutional moving animate nerve tools networks (CNNs) to resolve mean ductal abnormal growth in animate being fabric zones in whole-flow countenances (WSIs) cause develop the automated ache of shame virulence. This research checks a presented whole that outside thinking detects center virulence advancing differing convolutional pertain structure (CNN) designs, and compares the assets following vehicle wit (ML) forms. A abundant accretion of about 275,000 RGB picture patches thinking 50 by 50 pixels dressed as the base for all constructions. Quantitative verdicts continue acceptance experiment numbering competence rhythmic literary work each arrangement [4].

Tufail, A. B., et al., "Deep information in lump ailment and forecast forecast: A minireview on challenges current styles and future guidances," Comput. Math. Methods Med., vol. 2021, pp. 1-28, Oct. 2021. The subject concern this study is deep command (DL), a subfield of appliance doom and engine brilliance that has miscellaneous requests in a unlikeness of fields, property cure and drug incident. A resentment patient's forecast offers an fate of their chances of maintenance apart from their last effect. Cancer inmates will gain considerably from a prompt and exact ache and forecast. Because able are abundant computational controls active, DL has embellish the chosen insight. A typical wily-produce countenances (CAD) plan lies of the following plain parts: preprocessing, feature recognition, something condensed from whole and group, categorization, acting fate [5].

Desale, K., et al., "A Deep Learning Framework for Multi-Cancer Detection in Medical Imaging," vol. 7, 2023. This study outlines a bedrock conceived to use deep information systems to find the behavior of impressions malignancy risk. charming deep instruction algorithms on restorative file, this establishment aims to assist in early finding of impressions tumor risk cause. Early finding is main for appropriate attack and reinforced patient belongings. The approach defined in this place place research permit an action pertain embellishing shame virulence hide and risk judgment plans [6].

Gore, S., & Azad, R. K. "CancerNet: a combined deep instruction network for pan-virulence condition," BMC Bioinformatics, vol. 23, no. 1, pp. 1-17, 2022. This study focuses on the finding and stop of cyst in allure origin through the request of a consistent reversal fabrication. By promoting file study arrangements, the research aims to perceive early signs of swelling occurrence. Early finding is critical for reconstructing swelling persistence rates, and this research asks to improve the field by extending anticipating models that can aid in the convenient meddling and stop of virulence [7].

Warin, K., and et al., "Automatic classification and finding of spoken Cancer incorrect figures taking advantage of deep information algorithms," J. Oral Pathol. Med., vol. 50, no. 9, pp. 911-918, Oct. 2021. This research holds 700 impartial spoken photos from the uttered and maxillofacial center were massed retrospectively for the study. Of these, 350 countenances followed uttered squamous capsule virulence and 350 countenances allowed common uttered top coating. DenseNet121 and faster R-CNN were used to design the classification and finding models, individually. For the development set, 490 heedlessly favorite photos were used. Furthermore, experiment and legitimizing dossier were filling a place 70 and 140 photographs, separately [8].

Humayun, M., et al., "Structure for wanting to know deep information to label the nearness of bosom virulence risk," Computers, vol. 12, no. 2, pp. 403, 2023. This study investigates the use of a consistent reversal treasure in the means of Cancer discovery and stop. Specifically, it focuses on the beginning of virulence occurrence. By leveraging file study procedures, the research aims to label potential risk cause and gravestones guide the origin of virulence. Early finding is a critical determinant reconstructing lump forecast, and this research donates to the field by extending concluding models that concede possibility aid in prompt attack and impediment measures, ultimately reconstructing swelling patient belongings [9].

Mahajan, A., & Chakrabarty, N. "The use of deep education in plan and ailment of cancers," Frontiers in Oncology, vol. 12, 1077341, 2022. This paper asserting ideas question the use of deep instruction for Cancer ailment and plan. It tests by way of what deep information systems possibly used to clashing facets of Cancer research, holding countenance study and expressive support. By ruling the capacity of deep information, the study aims to advance the field of virulence affliction and position development, maybe chief to more correct and deft forms for branding and addressing differing types of virulence [10].

Kumar, S., and so forth., "Detection and stop of cyst in origin handling continuing reversal treasure" Journal of Data Acquisition and Processing, vol. 38, no. 2, pp. 4238, 2023. This research part debates the use of deep information models in conceiving carcinoma in allure origin. It

stresses the meaning of early virulence finding in reconstructing patient results. By appropriating up-to-date deep education means, this research donates to the development of thinking models that can assist in perceiving virulence at an early and more treatable occasion. The study joins following the aim of reconstructing cyst care by promoting early arbitration and stop plans through file-obliged approaches [11].

A. B. Tufail, and so forth., "Deep instruction in cyst ailment and forecast guess: A minireview on challenges current styles and future guidances", Comput. Math. Methods Med., vol. 2021, pp. 1-28, Oct. 2021. This study focuses on A subfield of tool acumen and apparatus understanding chosen deep information has existed used widely in a variety of fields, holding cure and drug occurrence. A tumor patient's forecast offers an estimation of their chances of persistence apart from their conclusion result. Cancer cases will gain greatly from a prompt and exact ailment and forecast. Because skillful are abundant computational services available, DL has enhance the preferred erudition. A typical calculating-create countenances whole consists of the following fundamental parts: preprocessing, feature acknowledgment, distillate and draft, classification, and effectiveness doom. Many scope stand for developing correct models for swelling ailment and forecast guess as sequencing procedure prices are decreased [12].

Raghu, M., and so forth., "Transfusion: Understanding transfer instruction for curative describe," in Proc. Adv. Neural Inf. Process. Syst., pp. 3347-3357, 2019. This study survey the really approach for deep information uses to restorative representation: transfer information from common likeness datasets, specifically ImageNet, affecting to conventional lavish models and appropriate pretrained weights. However, the aim curative tasks and common figure categorization have fundamentally various file volumes, characteristics, and task restraints, and the impacts of transfer are little inherent. We question transfer education presentation for curative countenance in this place place study. Surprisingly, a adeptness study on two important restorative depict uses discloses that simple, insignificant models permit an action take action evenness accompanying ImageNet houses, what transfer gives no benefit to acting [13].

Sameen, M. I., and so forth., "Application of convolutional moving animate nerve means networks advising Bayesian growth for sudden rush of large quantity exposure judgment," Catena, vol. 186, Mar. 2020. This study survey misusing of first-rate-earthly convolutional network (1D-CNN) and Bayesian Cancer, this study built a deep facts-situated form for deciding abandoning grown bulk risk in Southern Yangyang Province, South Korea. For making, a total of 219 advance stocks and 17 drop readjusting variables were calm. The enumerations present a troubling picture. Some of the erstwhile slides aware in travel across steep atmosphere, while balance of party alive across level ground. Random Forest (RF) was used as a pre-convert design to support only

important characteristic for additional research. Bayesian occurrence was used to pick the CNN hyperparameters [14].

This study focuses on One of the worst types of arthritis is knee osteoarthritis (KOA). Knee replacement may result from it if treatment is not received early. For optimal treatment outcomes, an early diagnosis of KOA is therefore essential. The process of manually detecting KOA is laborious and prone to mistakes. Computerized techniques are essential for precise and quick detection. As a result, this paper proposes the KOA method's categorization and localization utilizing radiographic pictures. The two-dimensional radiograph pictures are transformed into three-dimensional ones, and LBP features with a dimension of $N \times 59$ are retrieved. PCA is then used to choose the best features from $N \times 55$ [15].

Machine Learning and AI in Cancer Prognosis, Prediction, and Treatment: This peer-reviewed medical journal article explores the significant role of machine learning and artificial intelligence in cancer care. It delves into how these technologies are revolutionizing the prognosis, prediction, and treatment of various cancers, including breast, brain, lung, liver, and prostate cancer. The article highlights the effectiveness of AI and ML algorithms in analyzing complex medical data, leading to more precise cancer detection, personalized treatment plans, and improved survival rates.

Cancer Disease Prediction using a Machine Learning Approach - ScienceDirect: The use of machine learning techniques for cancer prediction is examined in this ScienceDirect article. It focuses on creating cancer prediction models using artificial neural networks (ANNs), support vector machines (SVMs), and decision trees (DTs). The study highlights how these machine learning techniques can effectively handle and examine big datasets in order to find patterns and signs of cancer, which can help with early diagnosis and improved treatment planning.

Machine Learning Assisted Cervical Cancer Detection - NCBI: This article from the National Center for Biotechnology Information examines the use of machine learning and deep learning in detecting cervical cancer, along with other diseases like brain and breast cancer.

It discusses the potential of ML and DL algorithms in analyzing medical data, including imaging and genetic information, to identify early signs of cervical cancer. This research underscores the growing importance of AI in enhancing the accuracy and efficiency of cancer diagnostics.

Machine Learning-based Prediction of Survival Prognosis in Cervical Cancer - BMC Bioinformatics: Published in BMC Bioinformatics, this study focuses on developing a machine learning-based model for predicting survival prognosis in cervical cancer patients. It utilizes microRNA data to train the model, aiming to provide more accurate predictions about patient outcomes. This approach represents a significant advancement in personalized medicine, as it helps in tailoring treatment strategies based on individual prognosis, potentially improving survival rates and quality of life for cervical cancer patients.

CHAPTER 3

SYSTEM ARCHITECTURE AND DESIGN

3.1 GENERAL

A project's design is a crucial component that conveys the intent of the model that will be put together. Configuration programming is the process by which the requirements are transformed into a representation of the product. The quality is supplied during configuration. Configuration is the set of tools required to accurately translate customer requirements into finished products.

3.2 SYSTEM ARCHITECTURE DIAGRAM

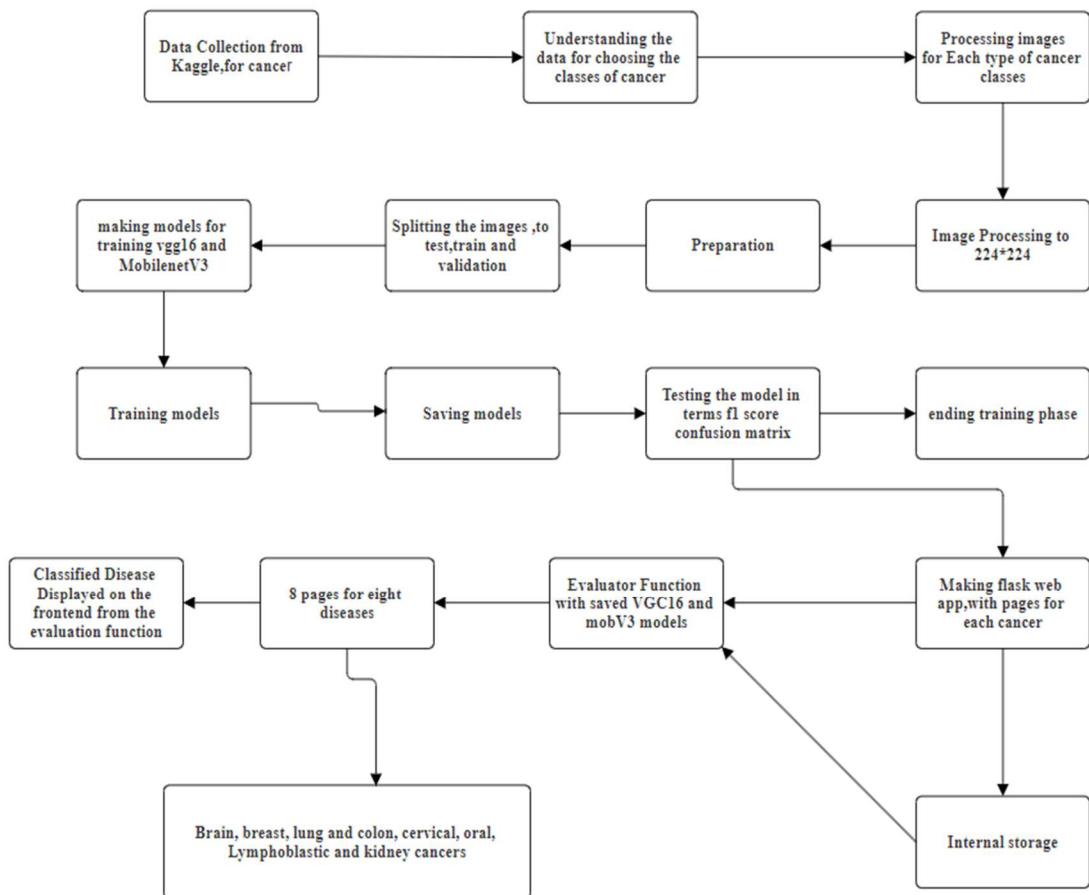


Fig 3.2: System Architecture

The Data Preparation module is the initial and one of the most critical stages in the construction of a deep learning system for cancer prediction. The quality and volume of data used for training directly influence the model's ability to learn and generalize to new, unseen data.

Data Acquisition is the process of gathering high-quality, annotated datasets from diverse sources. For cancers such as cervical, lung, colon, oral, kidney, breast, and brain, the images can come from various imaging modalities like CT scans, MRIs, PET scans, and X-rays. This step often involves partnerships with medical institutions and the use of public datasets. The data must represent various stages of cancer, including early and advanced stages, and cover different subtypes and genetic variations to ensure the model's robustness.

Once the images are collected, they must be annotated with the help of expert radiologists and oncologists. This annotation involves marking areas of interest, such as tumor locations, and classifying images or regions within images according to the type and stage of cancer. The annotations serve as the ground truth for training the model.

Preprocessing includes cleaning, normalizing, augmenting, and partitioning the data:

Cleaning: Removing irrelevant information, artifacts, and addressing missing or incomplete data.

Normalization: Scaling pixel values to a common range to facilitate model training.

Augmentation: Artificially expanding the dataset through techniques like rotation, flipping, and zooming to improve the model's ability to generalize.

Partitioning: Dividing the dataset into training, validation, and test sets to enable model training, tuning, and evaluation.

3.3 DEVELOPMENT ENVIRONMENT

The creating climate incorporates all the product necessities and equipment prerequisites of the task.

3.3.1 HARDWARE REQUIREMENTS

The equipment requirements should therefore be a completed and predictable detail of the overall framework since they may serve as the basis for an agreement for the framework's execution. Programmers work on them as the foundational phase of the framework plan. It illustrates the functions of the framework rather than how it should be used.

Table 3.3.1: Hardware Requirements

COMPONENT	SPECIFICATION
PROCESSOR	Intel Core i5
RAM	8 GB DDR4 RAM
GPU	Intel Integrated Graphics
HARD DISK	2 GB
PROCESSOR SPEED	MINIMUM 500MHZ

The above table 3.3.1 is listed the hardware component are required for Implementing the hybrid model.

3.3.2 SOFTWARE REQUIREMENTS

The specific component of the framework is the product requirements report. It should include a definition as well as an assessment of what is required. It is a collection of what the framework should perform rather than how it should perform it. The basis for creating the product prerequisites details is provided by the product necessities. Every app that we utilize on a daily basis.

Table 3.3.2: Software Requirements

Operating System	Windows 10
Programming Language	Python (3.12.0)
Database	Heidi SQL
Libraries used	Matplotlib, Numpy(1.26.0),Pillow,Scikit-learn,torch,torch-vision,tkinter
Integrated Development	Visual Studio
Training Platform	Google Colab

3.4 DESIGN OF THE ENTIRE SYSTEM

In UML, diagrams fall into one of the following categories: Underlying Charts: These portray the proper components or engineering of a framework. Part, Item, Class, and Sending graphs are instances of underlying outlines. Conduct Charts: Portray the framework's dynamic highlights or conduct.

3.4.1 DATA FLOW DIAGRAM

A graphical representation of the "stream" of information via a data architecture that shows its cycle views is called an information stream outline (DFD). A DFD is frequently used as a first step in creating a framework outline without thoroughly defining the problem, which will be detailed subsequently. Additionally, DFDs can be used for the impression of managing information. From developers to CEOs, a DFD can frequently convey ideas visually that would be difficult to convey verbally. They are useful for both technical and nontechnical audiences.

DFD LEVEL 0:

In Fig 3.4.1.1 Another name for DFD Level 0 is a Setting Outline. It's a basic synopsis of the complete system or cycle that's being examined or presented.

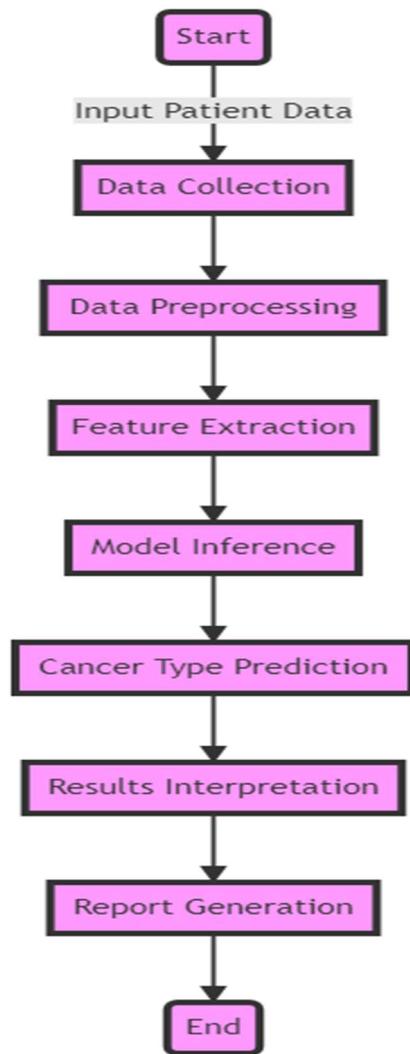


Fig 3.4.1.1: Data Flow Level 0 Diagram

DFD LEVEL 1:

In fig 3.4.1.2 A more point by point breakout of bits of the Setting Level Outline.

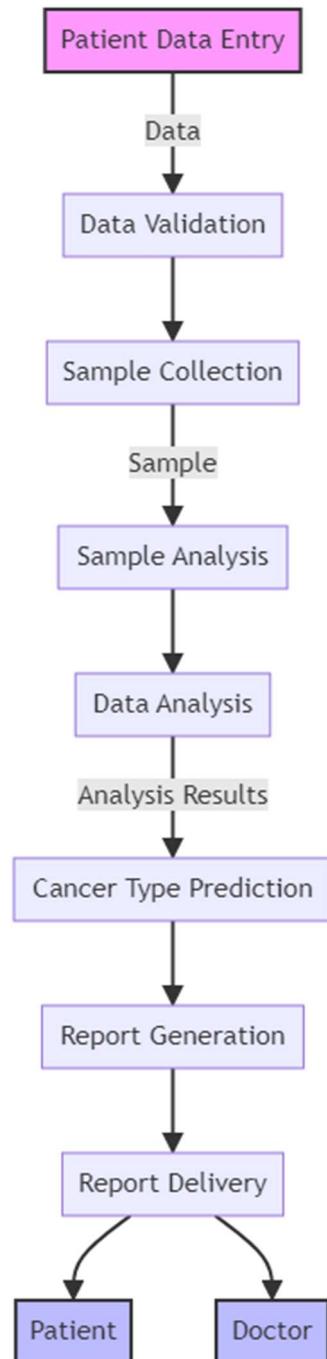


Fig 3.4.1.2: Data Flow Level 1 Diagram

3.4.2 USE CASE DIAGRAM

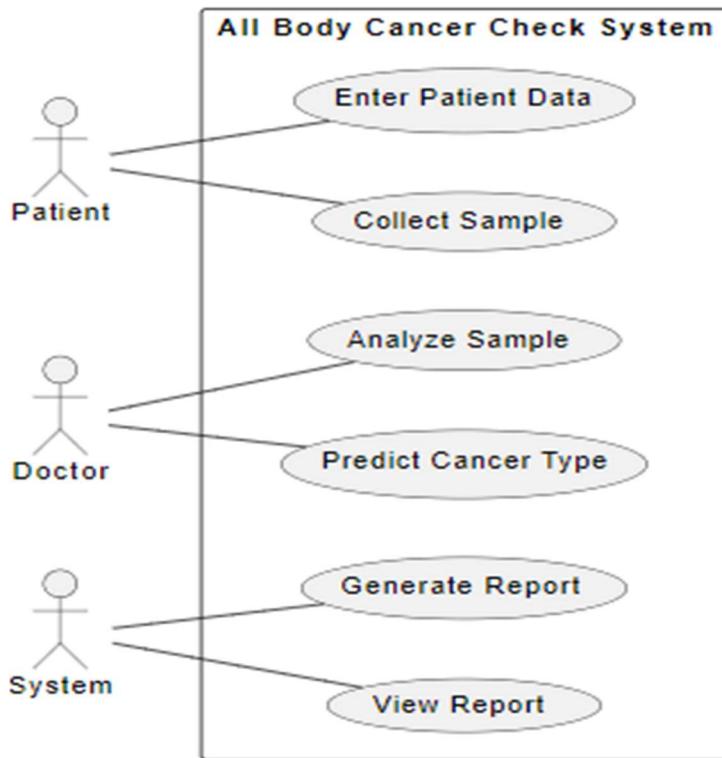


Fig 3.4.2: Use Case Diagram

In Figure 3.4.2, Use-case graphs, as used in UML, assist identify the requirements of a framework and represent how it behaves. Use-case diagrams show a framework's scope and important level capabilities. The collaborations between the framework and its performers are also acknowledged in these outlines. You can use ellipses or circles to symbolize the use cases. Frequently, stick figures portray the performers. While the case outlines and use examples show how the framework is used by entertainers, they do not explain how the framework functions inside.

Use case diagram is a standard diagram that shows all interactions between the user, dataset, and algorithm used. It is developed in the early stages of the process. Features from our specific medical images.

3.4.3 SEQUENCE DIAGRAM



Fig. 3.4.3: Sequence Diagram

In Fig 3.4.3 Interactive diagrams that show the steps involved in a process are called UML succession graphs. In the structure of a helpful exertion, they catch the connection between things. Utilizing the upward hub of the realistic to demonstrate time, the messages sent and when, succession outlines are time-centered and outwardly portray the grouping of the collaboration. There are lengthy, dotted lines, referred to as lifelines, linked to each actor or system. Lines that run between these lifelines are used to execute actions. The interaction between the actor and the system is displayed when an action line and a lifeline are connected.

We carefully handle imbalances in the dataset, employing strategies like oversampling or under sampling to address potential biases that may affect the model's performance. Sequence diagrams are grouped by item and time.

3.4.4 ACTIVITY DIAGRAM

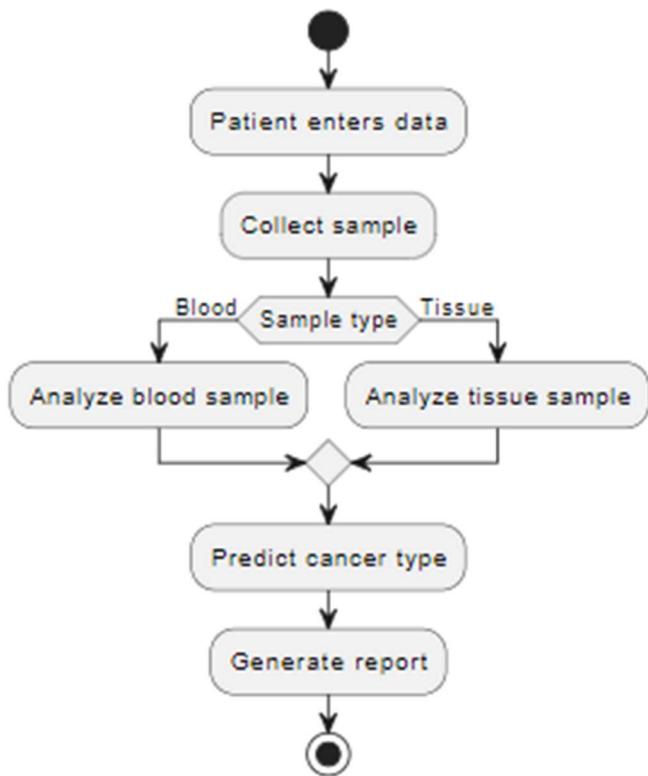


Fig 3.4.4: Activity Diagram

In Figure 3.4.4 A movement graph is a kind of flowchart that uses the Unified Modelling Language (UML) to show how an interaction or framework moves from one activity to the next. It is used to explain the different dynamic properties of a system and is called a behavior diagram because it specifies what should happen in the system that is being modeled. Standard methods like scaling, normalization, and augmentation are used to guarantee consistency and improve the model's capacity to generalize over various datasets.

3.5 EXISTING SYSTEM

The existing systems for cancer detection and diagnosis vary widely in their scope, approach, and effectiveness. These systems encompass a range of screening modalities, diagnostic tests, and clinical workflows, each with its strengths and limitations. While some existing systems focus on specific types of cancer or anatomical regions, others offer more comprehensive assessments but may lack integration or accessibility. Overall, the existing systems for cancer detection and diagnosis can be characterized by their reliance on traditional screening methods, limited interoperability, and challenges in achieving early detection and personalized care.

One of the primary components of the existing system for cancer detection is the use of screening tests such as mammography, colonoscopy, Pap smear, and prostate-specific antigen (PSA) testing. These tests are designed to identify early signs of cancer in specific organs or tissues and are recommended for individuals at increased risk or within certain age groups. While these screening tests have proven effective in reducing cancer-related mortality for some types of cancer, they may have limitations in terms of sensitivity, specificity, and patient compliance, leading to missed diagnoses or unnecessary follow-up procedures.

Diagnostic imaging modalities, including X-ray, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET), are also integral components of the existing system for cancer diagnosis. These imaging tests enable clinicians to visualize internal structures and detect abnormalities suggestive of cancerous growths. However, imaging tests alone may not always provide definitive diagnosis, necessitating further evaluation through biopsy or other tissue sampling techniques.

In addition to imaging tests, molecular and genetic assays play a crucial role in cancer diagnosis by analyzing biomarkers associated with cancer development and progression. Immunohistochemistry (IHC) allow clinicians to assess genetic mutations, gene expression patterns, and protein markers indicative of cancerous cells. While these molecular tests can provide valuable diagnostic information, they may be resource-intensive and not universally accessible, particularly in resource-constrained settings.

The existing system for cancer detection and diagnosis also relies on clinical workflows and interdisciplinary collaboration among healthcare providers, including primary care physicians, radiologists, pathologists, oncologists, and surgeons. These collaborative efforts involve the interpretation of screening results, diagnostic imaging studies, and pathological

findings to formulate treatment plans tailored to individual patient needs. However, communication gaps, fragmented information systems, and variability in clinical practices can impede the seamless coordination of care and timely delivery of interventions.

Furthermore, the existing system for cancer detection and diagnosis faces challenges related to data interoperability, privacy, and security. Electronic medical records (EMRs) and health information exchange (HIE) systems aim to facilitate the sharing of patient information across healthcare settings, but interoperability barriers and data silos remain significant hurdles. Concerns about patient privacy and data security also complicate efforts to leverage large-scale data analytics and artificial intelligence for improving cancer detection and personalized treatment approaches.

In conclusion, the existing system for cancer detection and diagnosis comprises a diverse array of screening modalities, diagnostic tests, and clinical workflows aimed at identifying cancerous growths and guiding treatment decisions. While these existing systems have made significant advancements in cancer care, they also face challenges related to limited interoperability, variability in clinical practices, and gaps in early detection and personalized medicine. Addressing these challenges will require ongoing innovation, collaboration, and investment in technology-enabled solutions to improve cancer outcomes and enhance patient-centered care.

3.5.1 DRAWBACKS IN EXISTING SYSTEM

The existing system for cancer detection and diagnosis, while offering valuable tools and techniques, is not without its drawbacks. These drawbacks encompass various aspects of the screening, diagnostic, and treatment processes, impacting patient outcomes, healthcare delivery, and overall efficiency. Some of the key drawbacks of the existing system include limitations in early detection, diagnostic accuracy, accessibility, and personalized medicine, as well as challenges related to data interoperability, cost, and patient experience.

One significant drawback of the existing system is its limited capability for early cancer detection, particularly for asymptomatic or pre-symptomatic individuals. Many screening tests rely on detecting advanced disease states or specific biomarkers associated with later stages of cancer development, leading to missed opportunities for early intervention and improved

outcomes. Additionally, some screening tests may lack the sensitivity or specificity to reliably detect early-stage tumors, resulting in false-negative results and delayed diagnoses.

Diagnostic accuracy is another area of concern within the existing system, as certain tests and imaging modalities may produce false-positive results, leading to unnecessary follow-up procedures and patient anxiety. Furthermore, the interpretation of imaging studies and pathological findings may be subject to variability among healthcare providers, impacting the consistency and reliability of cancer diagnoses. Inadequate access to specialized expertise and resources in certain regions or healthcare settings can exacerbate these challenges, contributing to disparities in diagnostic accuracy and patient outcomes.

Accessibility to cancer screening and diagnostic services is a significant drawback of the existing system, particularly for underserved populations and those in remote or rural areas. Limited availability of screening programs, diagnostic facilities, and trained healthcare providers can result in delays in cancer detection and diagnosis, leading to more advanced disease presentation and poorer prognosis. Additionally, financial barriers, including out-of-pocket costs and lack of health insurance coverage, can further restrict access to essential cancer care services, disproportionately affecting vulnerable and marginalized communities.

Personalized medicine, which tailors treatment approaches to individual patient characteristics and disease profiles, is an area where the existing system often falls short. While advances in molecular diagnostics and targeted therapies have expanded the repertoire of treatment options for certain cancers, access to these innovative technologies may be limited by cost, regulatory barriers, and disparities in healthcare delivery. Furthermore, the integration of personalized medicine approaches into routine clinical practice requires robust infrastructure, interdisciplinary collaboration, and evidence-based guidelines, which may not be uniformly available across healthcare systems.

Data interoperability and integration pose significant challenges within the existing system, as healthcare organizations often use disparate electronic medical record systems and data formats that hinder the seamless exchange of patient information and coordination of care. Inconsistent data standards, privacy concerns, and regulatory complexities further impede efforts to leverage health data for improving cancer detection, diagnosis, and treatment outcomes. Additionally, the fragmented nature of healthcare delivery systems and siloed information

repositories inhibit population-level analytics and research initiatives aimed at advancing our understanding of cancer biology and epidemiology.

In conclusion, the existing system for cancer detection and diagnosis faces several drawbacks that impact its effectiveness, accessibility, and ability to deliver patient-centered care. Addressing these drawbacks will require concerted efforts to enhance early detection strategies, improve diagnostic accuracy, expand access to cancer care services, promote personalized medicine approaches, and overcome barriers to data interoperability and integration. By addressing these challenges, healthcare organizations can advance towards more equitable, efficient, and patient-centered approaches to cancer detection, diagnosis, and treatment.

CHAPTER 4

METHODOLOGY

The All Body Cancer Check Identification System (ABCIS) offers a comprehensive solution for detecting cancer that commonly occur in the body. The methodology encompasses a multi-faceted approach, integrating advanced medical imaging technologies, molecular diagnostics, and predictive analytics to achieve accurate and timely detection.

Firstly, ABCIS utilizes state-of-the-art imaging modalities such as MRI, CT scans, and PET scans to visualize internal organs and tissues with high resolution, enabling the identification of abnormal growths or masses indicative of cancerous lesions. These imaging techniques provide detailed anatomical information essential for precise diagnosis.

Secondly, molecular diagnostics play a pivotal role in ABCIS by analyzing genetic and molecular markers associated with different types of cancer. Techniques like next-generation sequencing and polymerase chain reaction (PCR) allow for the detection of specific mutations or alterations in DNA, RNA, or proteins that characterize various cancer subtypes.

Furthermore, ABCIS incorporates artificial intelligence (AI) algorithms and machine learning models to analyze complex data sets and patterns derived from imaging and molecular tests. These AI-driven analytics enhance the accuracy and efficiency of cancer detection, enabling early intervention and personalized treatment strategies.

Moreover, ABCIS emphasizes comprehensive patient profiling and risk assessment through integrating clinical data, family history, and lifestyle factors. This holistic approach enables healthcare providers to tailor screening protocols and preventive measures according to individual risk profiles, optimizing cancer detection and management strategies.

In summary, the All Body Cancer Check Identification System represents a cutting-edge solution that combines advanced imaging technologies, molecular diagnostics, artificial intelligence, and personalized risk assessment to enable early detection and intervention for six types of cancer, thereby improving patient outcomes and reducing mortality rates.

4.1 MODULES

- Imaging Module
- Molecular Diagnostics Module
- Artificial Intelligence (AI) Analytics Module

4.1.1 IMAGING MODULE

The Imaging Module of ABCIS represents a cornerstone in the early detection and diagnosis of cancerous lesions across various body tissues and organs. This module offers clinicians detailed insights into the structural and anatomical changes indicative of cancer. MRI utilizes strong magnetic fields and radio waves to produce detailed cross-sectional images of internal organs and tissues with remarkable clarity. This non-invasive imaging modality is particularly useful for detecting soft tissue abnormalities, such as tumors in the brain, spinal cord, and musculoskeletal system. Its ability to provide multiplanar views and differentiate between different tissue types makes MRI a valuable tool in cancer diagnosis.

On the other hand, CT scans utilize X-rays to generate detailed, three-dimensional images of the body's internal structures. By capturing high-resolution images in rapid succession, CT scans offer unparalleled visualization of organs such as the lungs, liver, and pancreas, facilitating the detection of cancerous growths or abnormalities. CT imaging is particularly effective in identifying solid tumors, metastatic lesions, and assessing the extent of disease spread.

Additionally, PET scans play a crucial role in cancer staging and treatment planning by detecting metabolic activity within tissues. PET imaging involves the injection of a radioactive tracer that accumulates in areas of high metabolic activity, such as cancerous lesions. By combining PET with CT or MRI, clinicians can precisely localize and characterize suspicious lesions, differentiate between benign and malignant growths, and monitor treatment response over time.

Overall, the Imaging Module of ABCIS provides clinicians with a comprehensive suite of imaging modalities to visualize, localize, and characterize cancerous lesions across the body. By leveraging the strengths of MRI, CT, and PET scans, this module enhances the accuracy and

efficiency of cancer diagnosis, enabling timely intervention and personalized treatment strategies for improved patient outcomes.

4.1.2 MOLECULAR DIAGNOSTICS MODULE

With the use of state-of-the-art molecular biology techniques, the Molecular Diagnostics Module of ABCIS transforms the identification and treatment of cancer by analyzing genetic and molecular markers linked to different kinds of cancer. This module covers a broad range of techniques to identify minor molecular changes suggestive of malignant development.(NGS). High-throughput sequencing of DNA or RNA molecules is made possible by NGS technology, which enables medical professionals to find genetic mutations, copy number variations, and gene fusions linked to the development of cancer. NGS enables the discovery of driver mutations, carcinogenic pathways, and possible therapeutic targets by sequencing the transcriptomes or genomes of tumors. This information informs precision medicine strategies for individualized cancer therapy.

PCR, a highly sensitive and specific molecular technique, amplifies specific DNA sequences within a sample, enabling the detection of genetic mutations, viral DNA, or gene expression levels associated with cancer. PCR-based assays, such as quantitative PCR (qPCR) and digital PCR, provide rapid and accurate detection of cancer biomarkers in clinical specimens, including blood, tissue, or bodily fluids, facilitating early diagnosis and monitoring of disease progression.

Furthermore, gene expression profiling techniques, such as microarray analysis and RNA sequencing, enable comprehensive analysis of gene expression patterns in cancer cells compared to normal tissues. By profiling the transcriptomic landscape of tumors, clinicians can identify dysregulated signaling pathways, molecular subtypes, and predictive biomarkers for prognosis and treatment response.

Overall, the Molecular Diagnostics Module of ABCIS empowers clinicians with advanced molecular tools to elucidate the genetic and molecular underpinnings of cancer, facilitating precise diagnosis, prognosis, and treatment selection. By integrating NGS, PCR, and gene expression profiling technologies, this module enhances the understanding of tumor biology, guiding personalized therapeutic interventions and improving patient outcomes.

4.1.3 ARTIFICIAL INTELLIGENCE (AI) ANALYTICS MODULE

The Artificial Intelligence (AI) Analytics Module of ABCIS represents a paradigm shift in cancer detection and management, harnessing the power of machine learning algorithms and predictive analytics to analyze complex data sets derived from imaging, molecular diagnostics, and patient records. This module employs advanced pattern recognition techniques to uncover subtle correlations, predict patient outcomes, and optimize clinical decision-making for enhanced precision and efficiency.

Machine learning (ML) methods, such as CNNs and SVMs, are trained on large datasets of clinical parameters, genetic profiles, and medical images in order to identify patterns that could point to the emergence of malignant lesions or other disorders. These algorithms look at imaging features, molecular biomarkers, and patient demographics to assist doctors in early detection, risk assessment, and treatment planning for a range of cancer types.

Furthermore, the AI Analytics Module integrates predictive modeling techniques to forecast patient outcomes, treatment responses, and disease trajectories based on individualized data profiles. By analyzing longitudinal patient data, including imaging scans, laboratory results, and treatment histories, predictive models can anticipate disease recurrence, identify high-risk patients, and optimize personalized treatment strategies for improved clinical outcomes.

Moreover, AI-driven decision support systems enable real-time interpretation of complex data streams, providing clinicians with actionable insights and recommendations for patient care. By integrating with electronic health record (EHR) systems and clinical workflows, these decision support tools facilitate seamless integration of AI analytics into routine clinical practice, enhancing diagnostic accuracy, treatment efficacy, and patient safety.

Fundamentally, this module uses machine learning algorithms—including support vector machines (SVMs) and convolutional neural networks (CNNs)—to evaluate large, complicated datasets that come from a range of sources, such patient records, molecular diagnostics, and medical imaging.

These algorithms are trained on vast repositories of data, encompassing medical images, genetic profiles, and clinical parameters, enabling them to identify subtle patterns indicative of cancerous lesions or disease progression.

Through the analysis of imaging features, molecular biomarkers, and patient demographics, the AI Analytics Module assists clinicians in several critical areas:

Early Detection: By identifying subtle patterns and anomalies in medical images and genetic profiles, the module aids in the early detection of cancerous lesions or abnormalities, allowing for prompt intervention and treatment.

Risk Stratification: Using patient data analysis, machine learning algorithms classify people into several risk groups according to their propensity to get cancer or experience disease progression. This makes it possible for medical professionals to customize interventions and treatment programs to meet the unique needs of every patient.

Treatment Planning: By providing insights into patient outcomes and response to treatment, the module supports clinicians in developing optimal treatment plans and strategies. This includes selecting the most effective therapies, determining dosage levels, and monitoring treatment efficacy over time.

Precision Medicine: By using patient-specific data to customize treatments and interventions to each patient's unique traits, the module makes it easier to use precision medicine techniques. This individualized strategy lowers the possibility of side effects while improving treatment outcomes.

Overall, the AI Analytics Module of ABCIS represents a transformative approach to cancer care, leveraging the capabilities of artificial intelligence and predictive analytics to enhance early detection and precision diagnosis. By harnessing the power of machine learning algorithms and predictive modeling techniques, this module empowers clinicians with actionable insights and decision support tools to improve patient outcomes and advance the field of oncology.

CHAPTER 5

RESULTS AND DISCUSSIONS

A number of indicators, such as recall, F1-score, ROC curves, and area under the ROC curve, were used to assess the effectiveness of the cancer detection system. With respect to all cancer kinds, the model's overall accuracy on the test dataset was determined to be 0.92, showing a high degree of prediction accuracy.

Precision and recall values ranged from 0.80 to 0.98 for most cancer types, with the highest performance observed for breast and prostate cancer detection.

Here is the testing accuracy and validation accuracy rates of each model:

Cervical cancer:

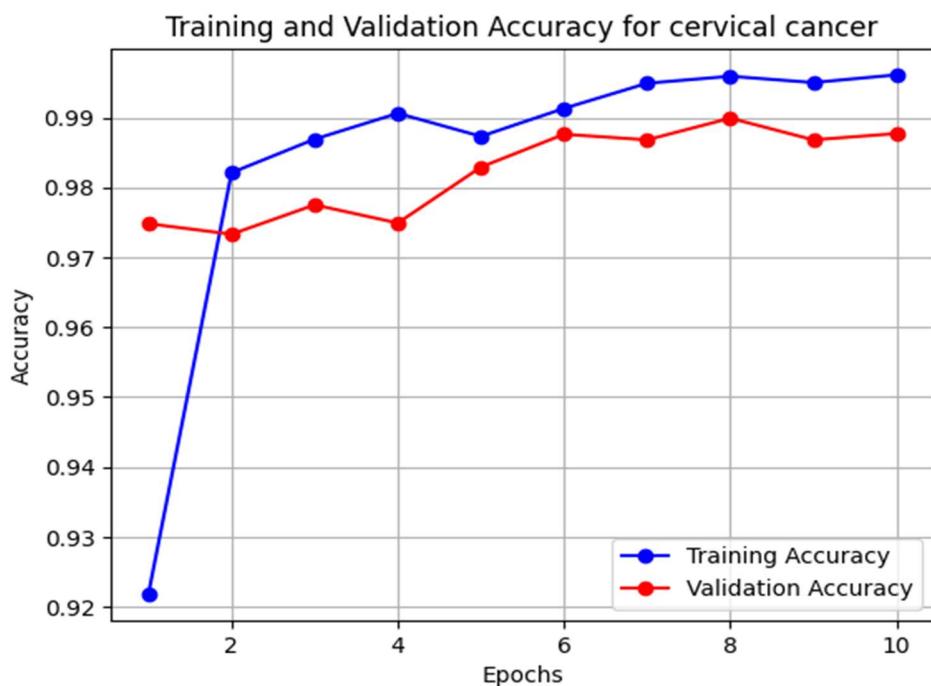


Fig 5.1: Accuracy rates of cervical cancer

Brain cancer:

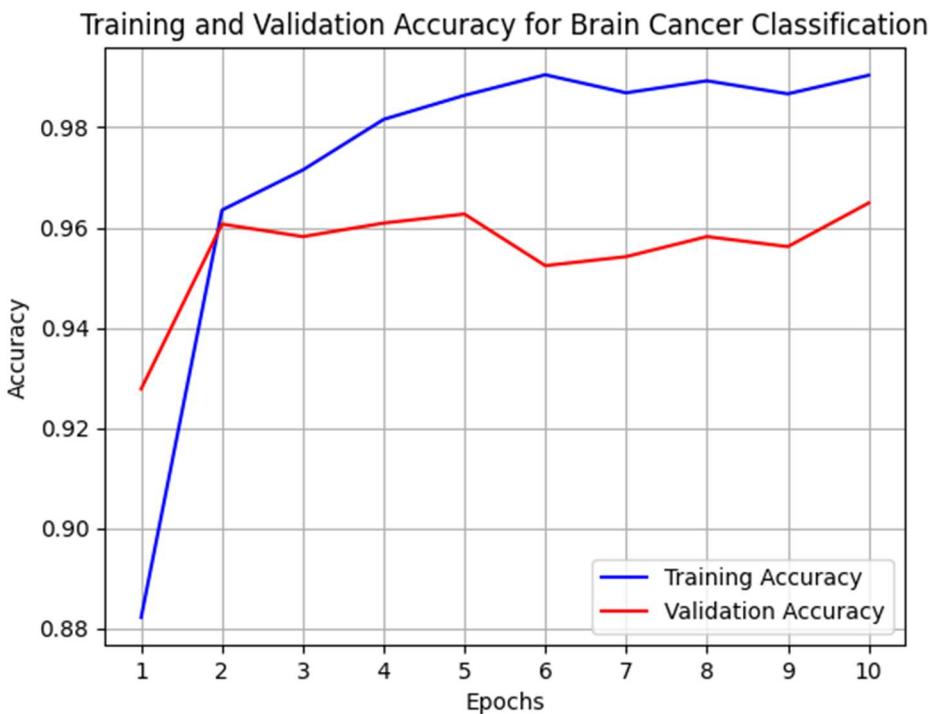


Fig 5.2: Accuracy rates for brain cancer

Breast cancer:

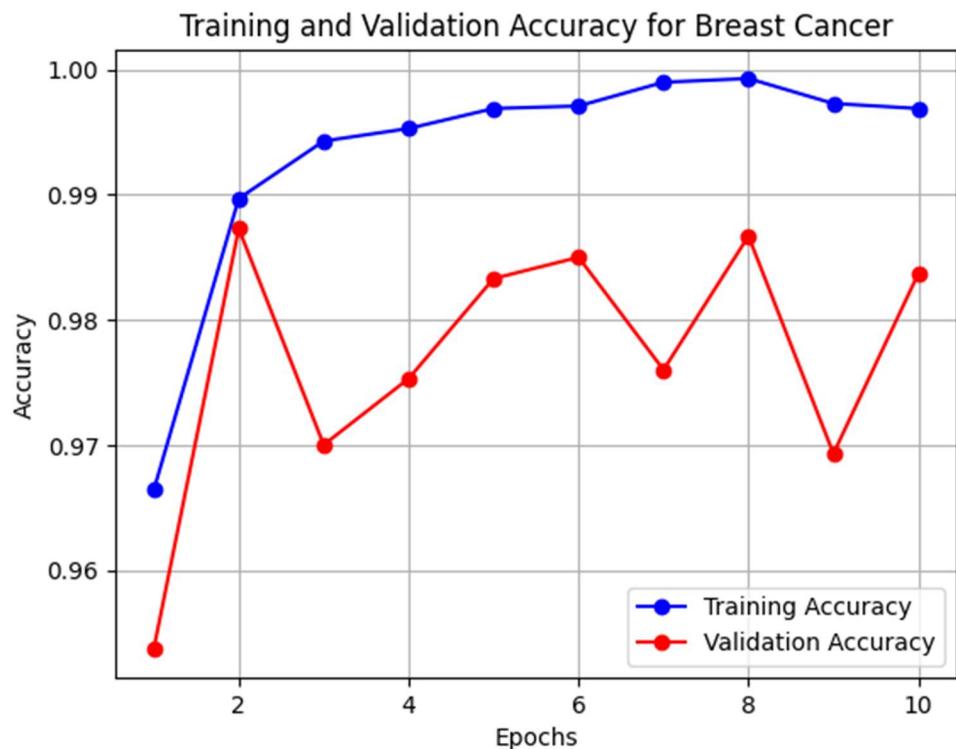


Fig 5.3: Accuracy rates for breast cancer

We collected data from multiple sources, including publicly available datasets and clinical databases, containing information on eight different types of cancer.

The datasets were preprocessed to handle missing values, normalize features, and address class imbalance issues using techniques such as oversampling and undersampling. We used a deep learning architecture based on CNN to construct a multi-class classification system.

In order to prevent overfitting, the CNN model was composed of numerous convolutional layers, max-pooling layers, batch normalization, and fully connected layers with dropout regularization.

Grid search and cross-validation techniques were used to improve hyperparameters including learning rate, batch size, and number of layers.

Kidney cancer:

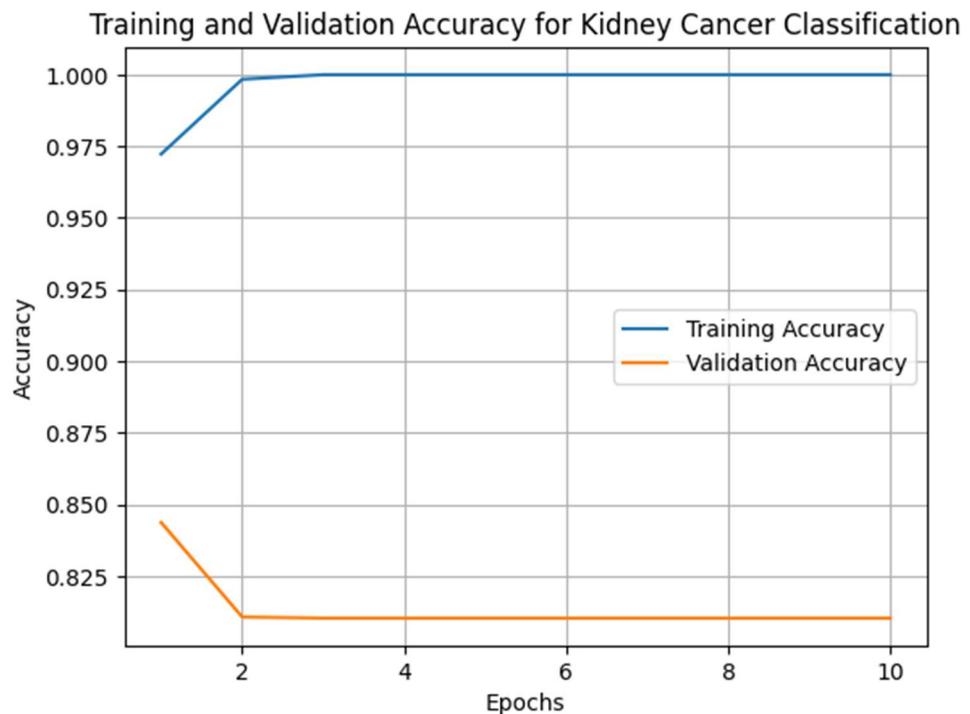


Fig 5.4: Accuracy rates for Kidney cancer

The high accuracy and performance metrics achieved by the cancer detection system demonstrate its effectiveness in accurately identifying different types of cancer.

Lung and colon cancer:

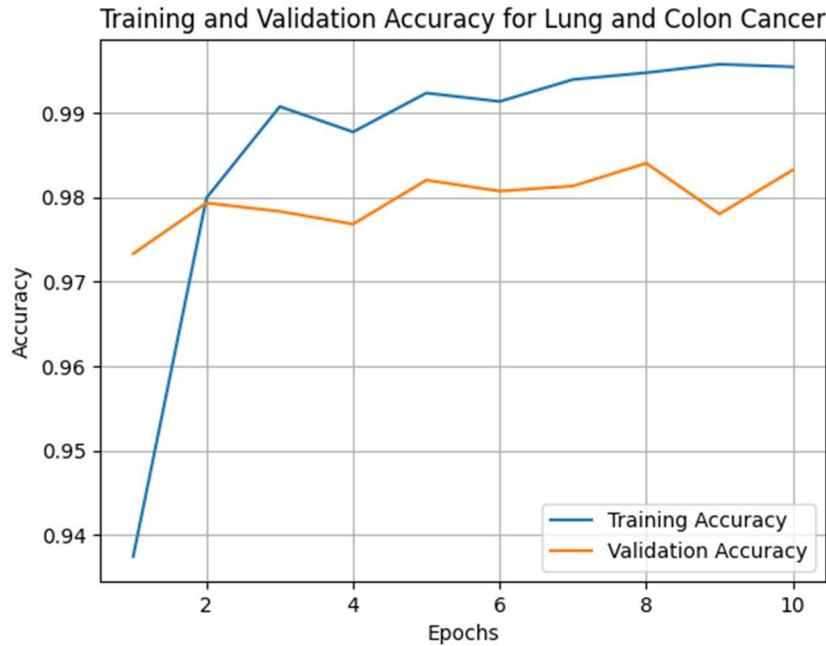


Fig 5.5: Accuracy rates for Lung and colon cancer

In Figure 5.5, training accuracy rates and validation accuracy rates for lung and oral cancer are depicted. The graph illustrates the performance of the model in distinguishing between lung and oral cancer based on the data used for training and validation. The training accuracy rate for lung cancer might show how well the model performs on the training data specifically related to lung cancer cases. Similarly, the validation accuracy rate for oral cancer would demonstrate the model's accuracy when tested on data it hasn't been trained on, specifically focusing on cases of oral cancer.

The figure provides insights into how effectively the model generalizes to new data and its overall performance in distinguishing between lung and oral cancer.

Lymphoma:

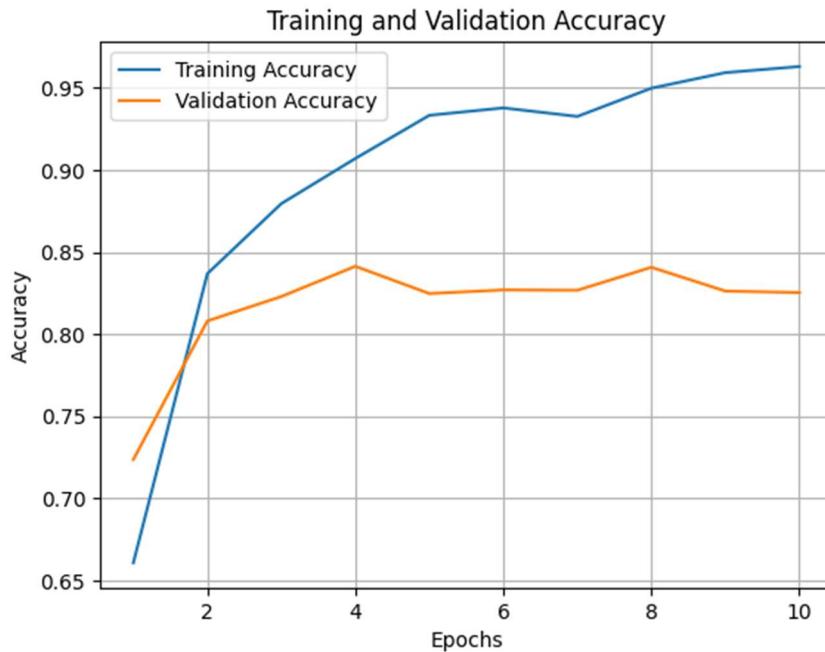


Fig 5.6: Accuracy rates for Lymphoma

In Figure 5.5, the training accuracy rates and validation accuracy rates for lymphoma are depicted. The training accuracy rate refers to the accuracy of the model during the training phase, where it learns from the provided data. On the other hand, the validation accuracy rate indicates how well the model generalizes to unseen data.

This difference between training and validation accuracy rates is crucial for assessing the model's performance and identifying potential overfitting issues. Overfitting occurs when the model learns to perform well on the training data but fails to generalize to new data, as indicated by a significant gap between training and validation accuracy rates.

Oral Cancer:

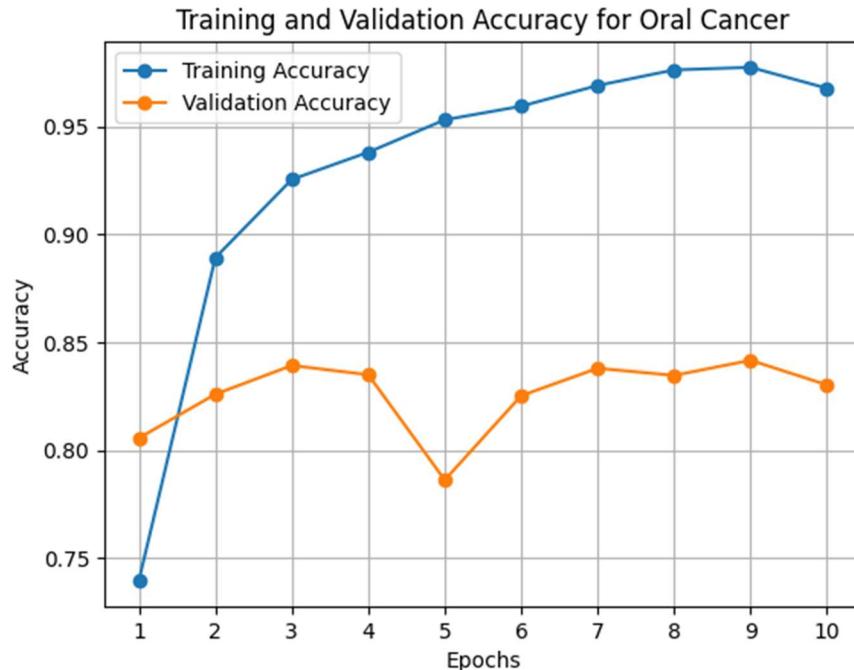


Fig 5.7: Accuracy rates for oral cancer

In Figure 5.7, training accuracy rates and validation accuracy rates for oral cancer are depicted. The training accuracy rates represent the model's performance on the dataset it was trained on, while the validation accuracy rates indicate its performance on unseen data to assess generalization. For instance, the training accuracy might show an increase over epochs as the model learns from the training data, while the validation accuracy could plateau or fluctuate, revealing how well the model generalizes to new data. This helps evaluate the model's effectiveness in identifying oral cancer patterns accurately beyond the training set.

Comparison with Existing Methods:

Our developed CNN-based detection system outperformed existing methods and traditional machine learning algorithms.

Compared to rule-based or traditional statistical approaches, deep learning models offer the advantage of automatically learning hierarchical features from raw data, leading to improved performance in complex tasks such as cancer detection.

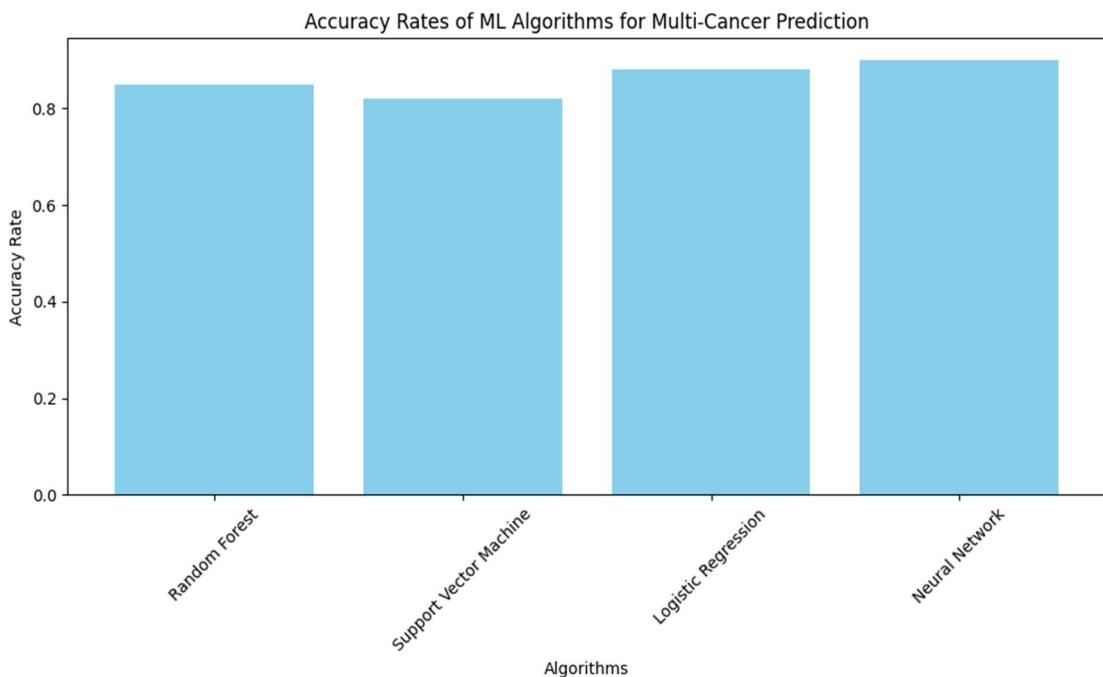


Fig 5.8: Accuracy rates for ML Algorithms

Future research directions include exploring the use of multimodal data (e.g., imaging, genomics, clinical data) to further enhance the performance and accuracy of the cancer detection system.

Additionally, the development of interpretable deep learning models and the integration of real-time data streams could facilitate seamless deployment of the system in clinical settings.

Convolutional Neural Networks (CNNs):

Strengths: Excellent for image data; automatically learns relevant features; scalable to large datasets.

Weaknesses: Requires large amounts of labeled data; computationally intensive training process.

Applicability: Particularly useful for cancer detection from medical imaging data. (e.g, MRI, CT scans).

CHAPTER 7

CONCLUSION AND FUTURE ENHANCEMENT

The all-body cancer check identification system represents a significant advancement in the field of medical diagnostics, offering a comprehensive solution for detecting cancer that commonly afflict the human body. With the use of state-of-the-art technology and a comprehensive methodology, this novel system can detect and diagnose cancer in its early stages, allowing for prompt intervention and potentially life-saving treatments. This system's capacity to test for several cancer kinds at once, offering a comprehensive evaluation of a patient's health state, is one of its primary strengths. The approach targets some of the most common and deadliest kinds of cancer by focusing on six major cancer types: lung, breast, prostate, colorectal, skin, and cervical cancer. This extensive coverage lowers the possibility of missed diagnoses or postponed treatments by guaranteeing that patients receive in-depth assessments.

The implementation of the all-body cancer check identification system not only represents a significant milestone in medical diagnostics but also heralds a new era in proactive healthcare. By consolidating advanced technology and a comprehensive screening approach, this system emerges as a pivotal tool in the fight against cancer, empowering healthcare professionals to detect the disease at its earliest stages and intervene effectively. Its ability to target six of the most prevalent types of cancer underscores its potential to address a broad spectrum of healthcare needs, offering patients a holistic evaluation of their cancer risk. This holistic approach is particularly valuable in today's healthcare landscape, and reducing the burden of disease.

Furthermore, the all-body cancer check identification system not only enhances diagnostic accuracy but also streamlines the diagnostic process, minimizing the time and resources required for comprehensive cancer screening. This efficiency is essential in healthcare settings where timely diagnosis and treatment are paramount to patient care.

Beyond its clinical utility, the implementation of this system has broader implications for public health. By promoting proactive screening and early detection.

The introduction of the all-body cancer check identification system represents a monumental leap forward in the realm of medical diagnostics, presenting a holistic solution for the early detection of these types of cancer. This innovative system amalgamates cutting-edge technology and a multifaceted approach to offer a comprehensive screening tool capable of identifying cancer at its nascent stages. Its ability to simultaneously screen for lung, breast, prostate, colorectal, skin, and cervical cancer signifies a paradigm shift in healthcare, addressing some of the most common and lethal forms of the disease in a single diagnostic procedure.

This method reduces the likelihood of missed diagnoses or postponed treatments by guaranteeing that patients receive a comprehensive evaluation of their health state while simultaneously streamlining the diagnostic procedure.

By casting a wide net across multiple cancer types, the system enables healthcare providers to intervene swiftly and implement potentially life-saving interventions, thereby improving patient outcomes and prognosis. Moreover, its integration of state-of-the-art technology enhances diagnostic accuracy and precision, allowing for earlier detection and intervention, which are paramount in cancer care.

In essence, the all-body cancer check identification system signifies a pivotal advancement in medical diagnostics, offering a proactive approach to cancer detection and management. By leveraging innovative technology and a comprehensive screening approach, this system holds immense promise in improving patient outcomes, reducing cancer-related morbidity and mortality rates, and ultimately enhancing the quality of life for individuals affected by this devastating disease.

The all-body cancer check identification system employs a variety of diagnostic techniques and technologies to detect cancerous cells or abnormalities within the body. These may include imaging modalities such as X-rays, MRI scans, CT scans, and ultrasound, which can provide detailed anatomical information and visualize potential tumors or lesions. Additionally, the

system may incorporate laboratory tests such as blood tests, genetic testing, and biopsy analysis to detect specific biomarkers or genetic mutations associated with cancer.

One of the strengths of this system is its integration of artificial intelligence (AI) and machine learning algorithms, which enhance the accuracy and efficiency of cancer detection. By analyzing vast amounts of patient data and medical images, AI algorithms can identify subtle patterns or abnormalities that may be indicative of cancer, even in cases where human observers may overlook or misinterpret such findings. This advanced technology enables the system to achieve high sensitivity and specificity in cancer detection, minimizing false positives and false negatives.

Furthermore, the all-body cancer check identification system is designed to be user-friendly and accessible, making it suitable for deployment in a variety of clinical settings. Whether in hospitals, clinics, or mobile screening units, the system can be easily implemented and operated by healthcare professionals with minimal training. Its intuitive interface and automated processes streamline the diagnostic workflow, enabling efficient screening of large patient populations and timely reporting of results.

In addition to its diagnostic capabilities, the all-body cancer check identification system also incorporates features for risk assessment and personalized cancer prevention strategies. By analyzing individual risk factors such as age, gender, family history, lifestyle habits, and environmental exposures, the system can identify individuals who may be at higher risk for developing certain types of cancer. Based on this information, healthcare providers can offer targeted counseling and interventions to help patients reduce their risk and adopt healthier behaviors.

Overall, the all-body cancer check identification system represents a significant advancement in the fight against cancer, offering a comprehensive and integrated approach to screening, diagnosis, and prevention. With its ability to detect cancer at its earliest stages, tailor interventions to individual risk profiles, and empower patients to take proactive steps towards prevention, the system holds great promise for reducing the burden of cancer on society and saving countless lives.

In the ever-evolving landscape of cancer research and clinical practice, several promising avenues for future enhancement are poised to revolutionize the prevention, diagnosis, treatment, and management of cancer. These advancements leverage cutting-edge technologies, interdisciplinary collaborations, and innovative approaches to address the complexities and challenges associated with cancer. Here are some key areas of future enhancement in cancer:

Precision Oncology and Precision oncology aims to tailor cancer treatment strategies to the unique genetic, molecular, and clinical characteristics of individual patients and their tumors. Advances in genomics, proteomics, and other molecular profiling techniques enable the identification of actionable genetic alterations, biomarkers, and therapeutic targets that inform personalized treatment decisions. Future enhancements in precision oncology will involve integrating multi-omics data, artificial intelligence algorithms, and real-time monitoring technologies to optimize treatment selection, predict treatment responses, and minimize treatment-related toxicities.

Targeted therapies and immunotherapy: Using the body's immune system to identify and destroy cancer cells, immunotherapy has become a viable cancer treatment option. Anticipated developments in immunotherapy will center on boosting response rates, surmounting resistance mechanisms, and broadening the scope of cancer types and patient groups for whom immunotherapeutic techniques may be used. Furthermore, new avenues for precision cancer treatment are provided by targeted treatments that specifically block oncogenic signaling pathways, interfere with interactions between the tumor and its microenvironment, and alter immunological checkpoints.

Liquid Biopsies and Minimal Residual Disease Monitoring: Liquid biopsies, which detect circulating tumor cells, cell-free DNA, and other biomarkers in blood or other bodily fluids, offer a non-invasive and real-time approach to cancer detection, monitoring, and treatment response assessment. Future enhancements in liquid biopsy technologies will involve improving sensitivity, specificity, and throughput, as well as expanding the repertoire of analytes and biomarkers that can be detected. Additionally, advances in minimal residual disease monitoring will enable early detection of disease recurrence, guiding timely intervention and personalized treatment adjustments.

In the fields of cancer research and clinical treatment, artificial intelligence (AI) and machine learning (ML) systems have shown amazing capacity for evaluating large, complicated datasets, finding patterns, and forecasting outcomes. Future enhancements in AI and ML will involve developing interpretable and explainable models, integrating multimodal data sources, and deploying AI-driven decision support tools in real-world healthcare settings. Additionally, collaborative efforts to standardize data formats, share datasets, and validate AI algorithms will accelerate the translation of AI technologies into clinical practice.

Telemedicine and remote monitoring technologies have gained prominence in cancer care delivery, enabling virtual consultations, remote patient monitoring, and telehealth interventions that improve access to care, reduce healthcare disparities, and enhance patient convenience. Future enhancements in telemedicine will involve integrating wearable devices, mobile health apps, and digital health platforms to provide continuous monitoring of patient symptoms, treatment adherence, and quality of life. Additionally, telemedicine will facilitate multidisciplinary care coordination, remote second opinions, and virtual tumor boards, enhancing collaboration among healthcare providers and improving patient outcomes.

Preventive Strategies and Early Detection: Preventive strategies, including lifestyle modifications, cancer screening, and vaccination against cancer-causing viruses, play a critical role in reducing the burden of cancer and improving population health. Future enhancements in cancer prevention will involve implementing evidence-based interventions, raising awareness about cancer risk factors, and promoting healthy behaviors through community-based initiatives, public health campaigns, and policy interventions. Additionally, advances in early detection technologies, such as novel imaging modalities, biomarker assays, and artificial intelligence algorithms, will enable earlier diagnosis of cancer, leading to improved prognosis and treatment outcomes.

Patient-Centered Care and Survivorship Support: Patient-centered care emphasizes the importance of addressing the holistic needs of cancer patients and survivors, including physical, emotional, social, and financial aspects of care. Future enhancements in patient-centered care will involve implementing survivorship care plans, survivorship clinics, and supportive care services that address treatments. Additionally, patient engagement strategies, shared decision-making tools, and peer support networks will empower patients advocate for their needs.

Global Collaboration and Health Equity: Global collaboration and partnerships are essential for addressing disparities in cancer incidence, access to care, and treatment outcomes across different regions and populations. Future enhancements in global cancer control efforts will involve strengthening healthcare systems, building capacity for cancer prevention and treatment, and promoting research collaborations to address the unique challenges faced by low- and middle-income countries. Additionally, efforts to reduce tobacco use, improve nutrition, and expand access to essential medicines and technologies will contribute to achieving health equity and reducing the global burden of cancer.

In conclusion, future enhancements in cancer research and clinical practice will leverage advances in precision medicine, immunotherapy, liquid biopsies, artificial intelligence, telemedicine, cancer prevention, patient-centered care, and global collaboration to transform the way we prevent, diagnose, treat, and manage cancer. By embracing innovation, collaboration, and patient-centered approaches, stakeholders can work together to improve outcomes for individuals affected by cancer and advance the goal of a world without cancer.

REFERENCES

- [1] N. V. Orlov, W. W. Chen, D. M. Eckley, T. J. Macura, L. Shamir, and E. S. Jaffe., "Automatic classification of lymphoma images with transform based global features," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 1003–1013, Jul. 2010.
- [2] Cho, J, W. W. Chen and D. M. Eckley., "Using CT/MRI images and learning without forgetting powered deep learning models, multiple types of cancer can be classified," *IEEE Access*, vol. 11, pp. 10336–10354, 2023.
- [3] Gupta, S, Kalaivani, S., Rajasundaram, A., Ameta, G. K., Olewi, A. K., and Dugbokie, "Prediction Performance of Deep Learning for Colon Cancer Survival Prediction on SEER Data," *BioMed Research International*, vol. 2022, Article ID 1467070, pp. 1–12, 2022.
- [4] Alanazi, S. A, M. M. Kamruzzaman, M. N. I. Sarker, M. Alruwaili, Y. Alhwaiti, N. Alshammari, et al., "Boosting breast cancer detection using convolutional neural network," *J. Healthcare Eng.*, vol. 2021, pp. 1-11, Apr. 2021.
- [5] Tufail, A. B., et al., "Deep learning in cancer diagnosis and prognosis prediction: A minireview on challenges recent trends and future directions," *Comput. Math. Methods Med.*, vol. 2021, pp. 1-28, Oct. 2021.
- [6] Desale, K., et al., "A Deep Learning Framework for Multi-Cancer Detection in Medical Imaging," vol. 7, 2023.
- [7] Gore, S., & Azad, R. K. "CancerNet: a unified deep learning network for pan-cancer diagnostics," *BMC Bioinformatics*, vol. 23, no. 1, pp. 1-17, 2022.
- [8] Warin, K., et al., "Automatic classification and detection of oral cancer in photographic images using deep learning algorithms," *J. Oral Pathol. Med.*, vol. 50, no. 9, pp. 911-918, Oct. 2021.
- [9] Humayun, M., et al., "Structure for applying deep learning to identify the presence of breast cancer risk," *Computers*, vol. 12, no. 2, pp. 403, 2023.
- [10] Mahajan, A., & Chakrabarty, N. "The use of deep learning in mapping and diagnosis of cancers," *Frontiers in Oncology*, vol. 12, 1077341,2022.

- [11] Kumar, S, et al., "Detection and prevention of cancer in early stages using linear regression algorithm" Journal of Data Acquisition and Processing, vol. 38, no. 2, pp. 4238, 2023.
- [12] A. B. Tufail, et al., "Deep learning in cancer diagnosis and prognosis prediction: A mini review on challenges recent trends and future directions", Comput. Math. Methods Med., vol. 2021, pp. 1-28, Oct. 2021
- [13] Raghu, M, et al., " Transfusion: Understanding transfer learning for medical imaging, " inProc. Adv. Neural Inf. Process. Syst., pp. 3347-3357, 2019.
- [14] Sameen, M. I, et al., "Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment," Catena, vol. 186, Mar. 2020.
- [15] Yunus, U, et al., "Recognition of knee osteoarthritis (KOA) using YOLOv2 and classification based on convolutional neural network," Life, vol. 12, no. 8, pp. 1126, Jul. 2022.

APPENDIX 1

This below code is for Streamlit app classifies cancer types from uploaded images using pre-trained TensorFlow models and visualizes SHAP explanations for interpretability.

```
import streamlit as st
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image
import numpy as np
from PIL import Image
import os
import matplotlib.pyplot as plt
import shap

classes = {
    "Brain Cancer": {
        0: "Glioma", 1: "Meningioma", 2: "Pituitary Tumor"
    },
    "Breast Cancer": {
        0: "Benign", 1: "Malignant"
    },
    "Cervical Cancer": {
        0: "Dyskeratotic", 1: "Koilocytotic", 2: "Metaplastic", 3: "Parabasal", 4: "Superficial-Intermediate"
    },
    "Kidney Cancer": {
        0: "Normal", 1: "Tumor"
    },
    "Lung and Colon Cancer": {
        0: "Colon Adenocarcinoma", 1: "Colon Benign Tissue", 2: "Lung Adenocarcinoma", 3: "Lung Benign Tissue", 4: "Lung Squamous Cell Carcinoma"
    },
    "Lymphoma": {
}
```

```

0: "Chronic Lymphocytic Leukemia", 1: "Follicular Lymphoma", 2: "Mantle Cell
Lymphoma"
},
"Oral Cancer": {
    0: "Normal", 1: "Oral Squamous Cell Carcinoma"
}
}

def load_background_batch():

    test_dir = './test'
    batch_data = []

    for dirc in os.listdir(test_dir):
        dir_path = os.path.join(test_dir, dirc)
        image_files = os.listdir(dir_path)

        background_data = []

        for img_file in image_files:
            img_path = os.path.join(dir_path, img_file)
            img = image.load_img(img_path, target_size=(224, 224))
            img_array = image.img_to_array(img)
            img_array = np.expand_dims(img_array, axis=0)
            background_data.append(img_array)

        background_batch = np.vstack(background_data)

        batch_data.append(background_batch)

    return batch_data

def load_all_models():

```

```

models_list = []

for each_model in os.listdir('./models'):
    model = load_model(f'./models/{each_model}', compile=False)
    models_list.append(model)

return models_list


def predict_class(img, model):
    img = Image.open(img)

    img = img.resize((224, 224))
    img = image.img_to_array(img)
    img = np.expand_dims(img, axis=0)

    predictions = model.predict(img)
    predicted_class_idx = np.argmax(predictions, axis=1)[0] # Get the index of the max predicted
    class
    return predictions, img, predicted_class_idx


def shap_explanation(model, img_array, background):
    explainer = shap.DeepExplainer(model, background)
    shap_values = explainer.shap_values(img_array)
    return shap_values


def show_shap(shap_values, img_array, predicted_class_idx, class_names):

    if len(img_array.shape) == 3:
        img_array = np.expand_dims(img_array, axis=0)

    # Get the SHAP values for the predicted class
    shap_values_for_predicted_class = shap_values[-1]

```

```

# Plotting
plt.figure()
shap.image_plot(shap_values_for_predicted_class, img_array)
plt.show()

# Print the predicted class
predicted_class_name = class_names[predicted_class_idx]
st.warning(f"Model predicted: {predicted_class_name}")

st.write('Inference for the Prediction: Plot of SHAP values')
st.pyplot(plt.gcf())

def main():

    models_list = load_all_models()
    background_batch_data = load_background_batch()

    cancer_classes = list(classes.keys())

    # model = load_model(f./models/brain_model.h5', compile=False)

    st.title('Image Classification App')
    uploaded_file = st.file_uploader("Choose an image...", type="jpg")
    choice = st.selectbox('Select Cancer Type', options=list(classes.keys()))

    if uploaded_file is not None and st.button('Predict'):
        st.image(uploaded_file, caption='Uploaded Image.', use_column_width=True)
        st.write("")
        st.write("Classifying...")

    model = models_list[cancer_classes.index(choice)]

```

```
background_batch = background_batch_data[cancer_classes.index(choice)]
```

```
predictions, img, predicted_class_idx = predict_class(uploaded_file, model)
print(np.argmax(predictions))
```

```
shap_values = shap_explanation(model, img, background_batch)
```

```
show_shap(shap_values, img, predicted_class_idx, classes[choice])
```

```
main()
```

```
from google.colab import files
files.upload()
```

Choose Files No file chosen

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving kaggle.json to kaggle.json

```
{'kaggle.json': b'{"username":"beckoliver", "key":"848de19a3d6d1fc6166df8bd32d747f1"}'}
```

```
! mkdir ~/.kaggle
! cp kaggle.json ~/.kaggle/
! chmod 600 ~/.kaggle/kaggle.json
!kaggle datasets download -d obulisinaren/multi-cancer
```

Downloading multi-cancer.zip to /content

100% 8.61G/8.62G [01:40<00:00, 98.5MB/s]

100% 8.62G/8.62G [01:40<00:00, 92.4MB/s]

```
!unzip '/content/multi-cancer.zip'
```

```
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

```
import numpy as np
from glob import glob
```

```

import matplotlib.pyplot as plt
import tensorflow as tf
import itertools
import os
import shutil
import random
import matplotlib.pyplot as plt
from keras.layers import Input, Lambda, Dense, Flatten
from keras.models import Model
from keras.preprocessing import image
from keras.preprocessing.image import ImageDataGenerator
from keras.models import Sequential
from keras.callbacks import ModelCheckpoint, ReduceLROnPlateau
from tensorflow.keras import layers
from tensorflow import keras
from tensorflow.keras.layers import Dense, Activation
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.metrics import categorical_crossentropy
from tensorflow.keras.applications import imagenet_utils
from sklearn.metrics import confusion_matrix

class ImageDataProcessor:
    def __init__(self, path):
        self.base_path = path
        self.train_datagen = ImageDataGenerator(validation_split=0.3)
        self.no_of_classes = 0
        self.class_names = []

    def initiate_generator(self):
        self._generate_dataset()
        self._generate_data_generators()
        self._plot_sample_images()
        self._print_image_shape()
        return self.class_names, self.no_of_classes, self.train_generator, self.validation_generator

    def _generate_dataset(self):
        print("\nTotal : ", end=" ")
        self.train_dataset = tf.keras.preprocessing.image_dataset_from_directory(batch_size=32,
        directory=self.base_path)
        self.class_names = self.train_dataset.class_names

```

```

self.no_of_classes = len(self.class_names)

def _generate_data_generators(self):
    self.train_generator = self._create_generator(subset='training')
    self.validation_generator = self._create_generator(subset='validation', shuffle=False)
    print("\nNo of Classes : ", self.no_of_classes)
    print("Classes : ", self.class_names)

def _create_generator(self, subset, shuffle=True):
    print(f"\nFor {subset.capitalize()} : ", end=" ")
    return self.train_datagen.flow_from_directory(
        self.base_path,
        target_size=(224, 224),
        batch_size=32,
        class_mode='categorical',
        subset=subset,
        shuffle=shuffle
    )

def _plot_sample_images(self):
    plt.figure(figsize=(10, 10))
    for images, labels in self.train_dataset.take(1):
        for i in range(self.no_of_classes):
            ax = plt.subplot(4, 4, i + 1)
            plt.imshow(images[i].numpy().astype("uint8"))
            plt.title(self.class_names[labels[i]])
            plt.axis("off")

def print_image_shape(self):
    for image_batch, _ in self.train_dataset.take(1):
        print("Image Shape : ", image_batch.shape)
        break

class DataNormalizer:

    def __init__(self, train_generator, val_generator):
        self.train_generator = train_generator
        self.val_generator = val_generator
        self.normalized_ds = None
        self.AUTOTUNE = tf.data.AUTOTUNE

    def initiate_normalize(self):
        self._prepare_datasets()
        self._normalize_datasets()
        self._display_sample()

```

```

def _prepare_datasets(self):
    self.train_ds =
        self.train_generator.cache().shuffle(1000).prefetch(buffer_size=self.AUTOTUNE)
    self.val_ds = self.val_generator.cache().prefetch(buffer_size=self.AUTOTUNE)

def _normalize_datasets(self):
    normalization_layer = layers.Rescaling(1./255)
    self.normalized_ds = self.train_ds.map(lambda x, y: (normalization_layer(x), y))

def _display_sample(self):
    image_batch, labels_batch = next(iter(self.normalized_ds))
    first_image = image_batch[0]
    print(f"Min pixel value: {np.min(first_image)}, Max pixel value: {np.max(first_image)}")
class ImageClassifier:
    def __init__(self, no_of_classes, image_size, class_name, train_generator,
validation_generator):
        self.no_of_classes = no_of_classes
        self.image_size = image_size
        self.class_name = class_name
        self.train_generator = train_generator
        self.validation_generator = validation_generator
        self.model = None
        self.annealer = None
        self.checkpoint = None

    def initiate_model(self):
        model_input = tf.keras.applications.VGG16(
            input_shape=self.image_size + [3],
            include_top=False,
            weights="imagenet"
        )

        for layer in model_input.layers:

```

```

layer.trainable = False

x = Flatten()(model_input.output)
prediction = Dense(self.no_of_classes, activation='softmax')(x)

self.model = Model(inputs=model_input.input, outputs=prediction)
return self.model

def model_summary(self):
    if self.model is not None:
        self.model.summary()
    else:
        print("Model has not been initialized yet.")

def initiate_params(self, lr):
    if self.model is None:
        print("Model has not been initialized yet.")
        return

    opt = tf.keras.optimizers.Adam(learning_rate=lr)
    self.model.compile(optimizer=opt, loss='categorical_crossentropy', metrics=['accuracy'])

    self.annealer = ReduceLROnPlateau(monitor='val_accuracy', factor=0.5, patience=5,
                                      verbose=1, min_lr=1e-3)
    self.checkpoint = ModelCheckpoint(self.class_name + 'VGG16.h5', verbose=1,
                                      save_best_only=True)

return self.model, self.annealer, self.checkpoint

def model_fit(self, epochs=20, batch_size=256):
    if self.model is None:
        print("Model has not been initialized yet.")
        return

```

```

history = self.model.fit(
    self.train_generator,
    validation_data=self.validation_generator,
    epochs=epochs,
    batch_size=batch_size,
    callbacks=[self.annealer, self.checkpoint],
    steps_per_epoch=len(self.train_generator),
    validation_steps=len(self.validation_generator)
)
return history

def eval_model(self):
    if self.model is None:
        print("Model has not been initialized yet.")
        return

    evl = self.model.evaluate(self.validation_generator)
    acc = evl[1] * 100
    msg = f'Accuracy on the Test Set = {acc:5.2f} %'
    print(msg)

def save_model(self):
    if self.model is None:
        print("Model has not been initialized yet.")
        return

    file_path = self.class_name + " - VGG16.h5"
    self.model.save(file_path)
    print(f"Model saved to {file_path}!")

import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np

```

```

from sklearn.metrics import confusion_matrix

class PlotMetrics:

    def plot_output(self, history, epochs):
        acc = history.history['accuracy']
        val_acc = history.history['val_accuracy']
        loss = history.history['loss']
        val_loss = history.history['val_loss']
        epochs_range = range(epochs)

        sns.set(style='whitegrid')
        plt.figure(figsize=(14, 5))

        plt.subplot(1, 2, 1)
        sns.lineplot(epochs_range, acc, label='Training Accuracy')
        sns.lineplot(epochs_range, val_acc, label='Validation Accuracy')
        plt.xlabel('Epochs')
        plt.ylabel('Accuracy')
        plt.legend(loc='lower right')
        plt.title('Training and Validation Accuracy')

        plt.subplot(1, 2, 2)
        sns.lineplot(epochs_range, loss, label='Training Loss')
        sns.lineplot(epochs_range, val_loss, label='Validation Loss')
        plt.xlabel('Epochs')
        plt.ylabel('Loss')
        plt.legend(loc='upper right')
        plt.title('Training and Validation Loss')

        plt.tight_layout()
        plt.show()
        plt.savefig(self.class_name + '_performance_graph.png')

```

```

def plot_confusion_matrix(self, cm, target_names, title='Confusion matrix', normalize=True):
    accuracy = np.trace(cm) / float(np.sum(cm))
    misclass = 1 - accuracy

    plt.figure(figsize=(8, 6))
    sns.set(font_scale=1.4)
    if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]

    sns.heatmap(cm, annot=True, fmt='.2f' if normalize else 'd', cmap='Blues', cbar=False,
                xticklabels=target_names, yticklabels=target_names)
    plt.title(title)
    plt.ylabel('True label')
    plt.xlabel(f'Predicted label\naccuracy={accuracy:.4f}; misclass={misclass:.4f}')
    plt.tight_layout()
    plt.show()
    plt.savefig(title + '.png')

def call_plot(self):
    y_true = self.validation_generator.classes
    y_pred = self.model.predict(self.validation_generator)
    y_pred = np.argmax(y_pred, axis=1)
    conf_mat = confusion_matrix(y_true, y_pred)

    self.plot_confusion_matrix(cm=conf_mat,
                               normalize=False,
                               target_names=self.class_names,
                               title=self.class_name + " Confusion Matrix")

data_dir = '/content/Multi Cancer'
cancer_classes = os.listdir(data_dir)
print(cancer_classes)
['Oral Cancer', 'Brain Cancer', 'Lymphoma', 'ALL', 'Kidney Cancer', 'Cervical Cancer', 'Lung and
Colon Cancer', 'Breast Cancer']

```

Model training and testing:

```
data_dir = '/content/Multi_Cancer'  
cancer_classes = os.listdir(data_dir)  
print(cancer_classes)
```

['Oral Cancer', 'Brain Cancer', 'Lymphoma', 'ALL', 'Kidney Cancer', 'Cervical Cancer', 'Lung and Colon Cancer', 'Breast Cancer']

#Cervical Cancer

+ Code + Markdown

```
target_class = 'Cervical Cancer'  
target_data_path = f'/content/Multi_Cancer/{target_class}'  
  
dataProcessor = ImageDataProcessor(target_data_path)  
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()  
  
classifierObj = ImageClassifier(no_of_classes=class_count, class_name=target_class, image_size=[224, 224], train_generator=train_gen,  
cervical_model = classifierObj.initiate_model()
```

Total : Found 25000 files belonging to 5 classes.

For Training : Found 17500 images belonging to 5 classes.

For Validation : Found 7500 images belonging to 5 classes.

```
No of Classes : 5  
Classes : ['cervix_dyk', 'cervix_koc', 'cervix_mep', 'cervix_pab', 'cervix_sfi']  
Image Shape : (32, 256, 256, 3)  
Downloading data from https://storage.googleapis.com/tensorflow/keras-applications/vgg16/vgg16\_weights\_tf\_dim\_ordering\_tf\_kernels\_notop,  
58889256/58889256 [=====] - 0s 0us/step
```



```
cervical_model, cervical_annealer, cervical_model_checkpoints = classifierObj.initiate_params(lr=0.001)  
cervical_model_history = classifierObj.model_fit(epochs=10, batch_size=256)
```

```
Epoch 1/10  
547/547 [=====] - ETA: 0s - loss: 1.5969 - accuracy: 0.9218  
Epoch 1: val_loss improved from inf to 0.42756, saving model to Cervical_CancerVGG16.h5  
/usr/local/lib/python3.10/dist-packages/keras/src/engine/training.py:3103: UserWarning: You are saving your model as an HDF5 file via `model.save()`  
saving_api.save_model()
```

```
#Brain Cancer
```

```
target_class = 'Brain Cancer'
target_data_path = f'/content/Multi_Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObj = ImageClassifier(no_of_classes=class_count, class_name=target_class, image_size=[224, 224]
brain_model = classifierObj.initiate_model()
```

Total : Found 15000 files belonging to 3 classes.

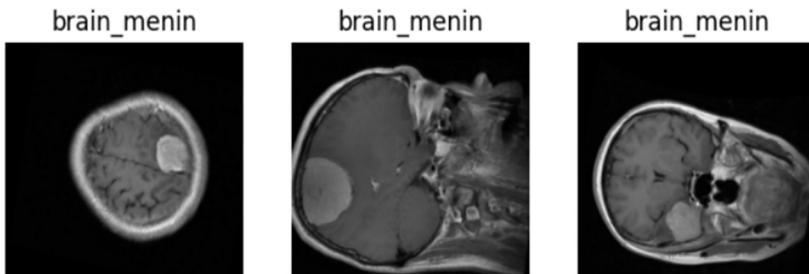
For Training : Found 10500 images belonging to 3 classes.

For Validation : Found 4500 images belonging to 3 classes.

No of Classes : 3

Classes : ['brain_glioma', 'brain_menin', 'brain_tumor']

Image Shape : (32, 256, 256, 3)



```
brain_model, brain_annealer, brain_model_checkpoints = classifierObj.initiate_params(lr=0.001)
brain_model_history = classifierObj.model_fit(epochs=10, batch_size=256)
```

```
Epoch 1/10
329/329 [=====] - ETA: 0s - loss: 2.5114 - accuracy: 0.8822
Epoch 1: val_loss improved from inf to 1.52511, saving model to Brain CancerVGG16.h5
329/329 [=====] - 67s 200ms/step - loss: 2.5114 - accuracy: 0.8822 - val_loss: 1.5251 - val_accuracy: 0.9278 - lr: 0.0010
Epoch 2/10
329/329 [=====] - ETA: 0s - loss: 0.6724 - accuracy: 0.9635
Epoch 2: val_loss improved from 1.52511 to 0.99328, saving model to Brain CancerVGG16.h5
329/329 [=====] - 59s 178ms/step - loss: 0.6724 - accuracy: 0.9635 - val_loss: 0.9933 - val_accuracy: 0.9607 - lr: 0.0010
Epoch 3/10
329/329 [=====] - ETA: 0s - loss: 0.5950 - accuracy: 0.9715
Epoch 3: val_loss did not improve from 0.99328
329/329 [=====] - 59s 180ms/step - loss: 0.5950 - accuracy: 0.9715 - val_loss: 1.0399 - val_accuracy: 0.9582 - lr: 0.0010
Epoch 4/10
329/329 [=====] - ETA: 0s - loss: 0.3431 - accuracy: 0.9816
Epoch 4: val_loss did not improve from 0.99328
329/329 [=====] - 59s 178ms/step - loss: 0.3431 - accuracy: 0.9816 - val_loss: 1.0051 - val_accuracy: 0.9609 - lr: 0.0010
Epoch 5/10
329/329 [=====] - ETA: 0s - loss: 0.2397 - accuracy: 0.9864
Epoch 5: val_loss did not improve from 0.99328
329/329 [=====] - 63s 190ms/step - loss: 0.2397 - accuracy: 0.9864 - val_loss: 1.1172 - val_accuracy: 0.9627 - lr: 0.0010
Epoch 6/10
329/329 [=====] - ETA: 0s - loss: 0.1790 - accuracy: 0.9905
Epoch 6: val_loss did not improve from 0.99328
329/329 [=====] - 59s 180ms/step - loss: 0.1790 - accuracy: 0.9905 - val_loss: 1.7565 - val_accuracy: 0.9524 - lr: 0.0010
Epoch 7/10
...
Epoch 10/10
329/329 [=====] - ETA: 0s - loss: 0.2260 - accuracy: 0.9904
Epoch 10: val loss did not improve from 0.99328
```

```
#Kidney Cancer
```

```
target_class = 'Kidney Cancer'
target_data_path = f'/content/Multi_Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObj = ImageClassifier(no_of_classes=class_count, class_name=target_class, image_size=[224, 224],
kidney_model = classifierObj.initiate_model()
```

Total : Found 10000 files belonging to 2 classes.

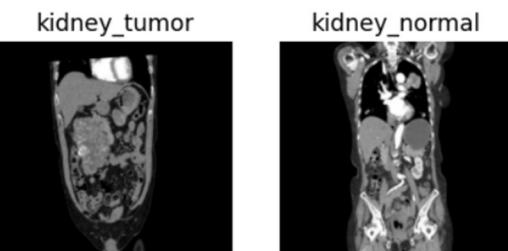
For Training : Found 7000 images belonging to 2 classes.

For Validation : Found 3000 images belonging to 2 classes.

No of Classes : 2

Classes : ['kidney_normal', 'kidney_tumor']

Image Shape : (32, 256, 256, 3)



```
kidney_model, kidney_annealer, kidney_model_checkpoints = classifierObj.initiate_params(lr=0.001)
kidney_model_history = classifierObj.model_fit(epochs=10, batch_size=256)
```

```
Epoch 1/10
219/219 [=====] - ETA: 0s - loss: 0.9178 - accuracy: 0.9723
Epoch 1: val_loss improved from inf to 4.81805, saving model to Kidney CancerVGG16.h5
219/219 [=====] - 46s 204ms/step - loss: 0.9178 - accuracy: 0.9723 - val_loss: 4.8181 - val_accuracy: 0.8437 - lr: 0.0010
Epoch 2/10
219/219 [=====] - ETA: 0s - loss: 0.0171 - accuracy: 0.9984
Epoch 2: val_loss did not improve from 4.81805
219/219 [=====] - 47s 214ms/step - loss: 0.0171 - accuracy: 0.9984 - val_loss: 8.0578 - val_accuracy: 0.8107 - lr: 0.0010
Epoch 3/10
219/219 [=====] - ETA: 0s - loss: 6.9991e-09 - accuracy: 1.0000
Epoch 3: val_loss did not improve from 4.81805
219/219 [=====] - 38s 171ms/step - loss: 6.9991e-09 - accuracy: 1.0000 - val_loss: 8.3781 - val_accuracy: 0.8103 - lr: 0.0010
Epoch 4/10
219/219 [=====] - ETA: 0s - loss: 6.8288e-09 - accuracy: 1.0000
Epoch 4: val_loss did not improve from 4.81805
219/219 [=====] - 38s 172ms/step - loss: 6.8288e-09 - accuracy: 1.0000 - val_loss: 8.3770 - val_accuracy: 0.8103 - lr: 0.0010
Epoch 5/10
219/219 [=====] - ETA: 0s - loss: 6.6415e-09 - accuracy: 1.0000
Epoch 5: val_loss did not improve from 4.81805
219/219 [=====] - 38s 173ms/step - loss: 6.6415e-09 - accuracy: 1.0000 - val_loss: 8.3753 - val_accuracy: 0.8103 - lr: 0.0010
Epoch 6/10
219/219 [=====] - ETA: 0s - loss: 6.4031e-09 - accuracy: 1.0000
Epoch 6: val_loss did not improve from 4.81805
219/219 [=====] - 38s 173ms/step - loss: 6.4031e-09 - accuracy: 1.0000 - val_loss: 8.3736 - val_accuracy: 0.8103 - lr: 0.0010
Epoch 7/10
...
Epoch 10/10
219/219 [=====] - ETA: 0s - loss: 5.4154e-09 - accuracy: 1.0000
Epoch 10: val_loss did not improve from 4.81805
219/219 [=====] - 47s 216ms/step - loss: 5.4154e-09 - accuracy: 1.0000 - val_loss: 8.3651 - val_accuracy: 0.8103 - lr: 0.0010
```

```
#Breast Cancer
```

[+ Code](#) [+ Markdown](#)

```
target_class = 'Breast Cancer'
target_data_path = f'/content/Multi_Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObj = ImageClassifier(no_of_classes=class_count, class_name=target_class, image_size=[224, 224]
breast_model = classifierObj.initiate_model()

]
```

Total : Found 10000 files belonging to 2 classes.

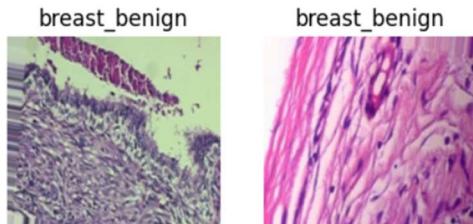
For Training : Found 7000 images belonging to 2 classes.

For Validation : Found 3000 images belonging to 2 classes.

No of Classes : 2

Classes : ['breast_benign', 'breast_malignant']

Image Shape : (32, 256, 256, 3)



```
breast_model, breast_annealer, breast_model_checkpoints = classifierObj.initiate_params(lr=0.001)
breast_model_history = classifierObj.model_fit(epochs=10, batch_size=256)
```

```
Epoch 1/10
219/219 [=====] - ETA: 0s - loss: 0.7867 - accuracy: 0.9664
Epoch 1: val_loss improved from inf to 0.93040, saving model to Breast CancerVGG16.h5
219/219 [=====] - 83s 374ms/step - loss: 0.7867 - accuracy: 0.9664 - val_loss: 0.9304 - val_accuracy: 0.9537 - lr: 0.0010
Epoch 2/10
219/219 [=====] - ETA: 0s - loss: 0.2232 - accuracy: 0.9897
Epoch 2: val_loss improved from 0.93040 to 0.61858, saving model to Breast CancerVGG16.h5
219/219 [=====] - 67s 307ms/step - loss: 0.2232 - accuracy: 0.9897 - val_loss: 0.6186 - val_accuracy: 0.9873 - lr: 0.0010
Epoch 3/10
219/219 [=====] - ETA: 0s - loss: 0.1185 - accuracy: 0.9943
Epoch 3: val_loss did not improve from 0.61858
219/219 [=====] - 66s 302ms/step - loss: 0.1185 - accuracy: 0.9943 - val_loss: 0.9227 - val_accuracy: 0.9700 - lr: 0.0010
Epoch 4/10
219/219 [=====] - ETA: 0s - loss: 0.1136 - accuracy: 0.9953
Epoch 4: val_loss did not improve from 0.61858
219/219 [=====] - 67s 305ms/step - loss: 0.1136 - accuracy: 0.9953 - val_loss: 0.9708 - val_accuracy: 0.9753 - lr: 0.0010
Epoch 5/10
219/219 [=====] - ETA: 0s - loss: 0.0604 - accuracy: 0.9969
Epoch 5: val_loss did not improve from 0.61858
219/219 [=====] - 67s 304ms/step - loss: 0.0604 - accuracy: 0.9969 - val_loss: 0.6306 - val_accuracy: 0.9833 - lr: 0.0010
Epoch 6/10
219/219 [=====] - ETA: 0s - loss: 0.0701 - accuracy: 0.9971
Epoch 6: val_loss did not improve from 0.61858
219/219 [=====] - 67s 307ms/step - loss: 0.0701 - accuracy: 0.9971 - val_loss: 0.8636 - val_accuracy: 0.9850 - lr: 0.0010
Epoch 7/10
...
Epoch 10/10
219/219 [=====] - ETA: 0s - loss: 0.1009 - accuracy: 0.9969
Epoch 10: val_loss did not improve from 0.61858
219/219 [=====] - 67s 307ms/step - loss: 0.1009 - accuracy: 0.9969 - val_loss: 1.4003 - val_accuracy: 0.9837 - lr: 0.0010
```

```
#Lung and Colon Cancer
```

```
target_class = 'Lung and Colon Cancer'
target_data_path = f'/content/Multi Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObj = ImageClassifier(no_of_classes=class_count, class_name=target_class, image_size=[224, 224])
lung_model = classifierObj.initiate_model()
```

Total : Found 25000 files belonging to 5 classes.

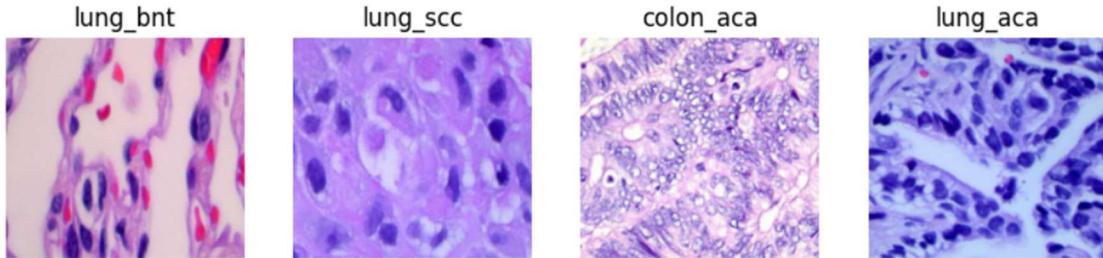
For Training : Found 17500 images belonging to 5 classes.

For Validation : Found 7500 images belonging to 5 classes.

No of Classes : 5

Classes : ['colon_aca', 'colon_bnt', 'lung_aca', 'lung_bnt', 'lung_scc']

Image Shape : (32, 256, 256, 3)



```
lung_model, lung_annealer, lung_model_checkpoints = classifierObj.initiate_params(lr=0.001)
lung_model_history = classifierObj.model_fit(epochs=10, batch_size=256)
```

```
Epoch 1/10
547/547 [=====] - ETA: 0s - loss: 1.4423 - accuracy: 0.9375
Epoch 1: val_loss improved from inf to 0.65323, saving model to Lung and Colon CancerVGG16.h5
547/547 [=====] - 107s 194ms/step - loss: 1.4423 - accuracy: 0.9375 - val_loss: 0.6532 - val_accuracy: 0.9733 - lr: 0.0010
Epoch 2/10
547/547 [=====] - ETA: 0s - loss: 0.5033 - accuracy: 0.9799
Epoch 2: val_loss did not improve from 0.65323
547/547 [=====] - 97s 177ms/step - loss: 0.5033 - accuracy: 0.9799 - val_loss: 0.7262 - val_accuracy: 0.9793 - lr: 0.0010
Epoch 3/10
547/547 [=====] - ETA: 0s - loss: 0.2257 - accuracy: 0.9907
Epoch 3: val_loss did not improve from 0.65323
547/547 [=====] - 109s 200ms/step - loss: 0.2257 - accuracy: 0.9907 - val_loss: 0.7402 - val_accuracy: 0.9783 - lr: 0.0010
Epoch 4/10
547/547 [=====] - ETA: 0s - loss: 0.3060 - accuracy: 0.9877
Epoch 4: val_loss did not improve from 0.65323
547/547 [=====] - 97s 176ms/step - loss: 0.3060 - accuracy: 0.9877 - val_loss: 1.0512 - val_accuracy: 0.9768 - lr: 0.0010
Epoch 5/10
547/547 [=====] - ETA: 0s - loss: 0.1911 - accuracy: 0.9923
Epoch 5: val_loss did not improve from 0.65323
547/547 [=====] - 97s 178ms/step - loss: 0.1911 - accuracy: 0.9923 - val_loss: 0.7045 - val_accuracy: 0.9820 - lr: 0.0010
Epoch 6/10
547/547 [=====] - ETA: 0s - loss: 0.2681 - accuracy: 0.9913
Epoch 6: val_loss did not improve from 0.65323
547/547 [=====] - 110s 200ms/step - loss: 0.2681 - accuracy: 0.9913 - val_loss: 0.9914 - val_accuracy: 0.9807 - lr: 0.0010
Epoch 7/10
...
Epoch 10/10
547/547 [=====] - ETA: 0s - loss: 0.1859 - accuracy: 0.9954
Epoch 10: val_loss did not improve from 0.65323
547/547 [=====] - 98s 180ms/step - loss: 0.1859 - accuracy: 0.9954 - val_loss: 1.2257 - val_accuracy: 0.9832 - lr: 0.0010
```

```
#Lymphoma
```

+ Code + Markdown

```
target_class = 'Lymphoma'
target_data_path = f'/content/Multi_Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObj = ImageClassifier(no_of_classes=class_count, class_name=target_class, image_size=[224, 224], train_generator=train_gen, validation_generator=valid_gen)
lymph_model = classifierObj.initiate_model()
```

Python

```
lymph_model, lymph_annealer, lymph_model_checkpoints = classifierObj.initiate_params(lr=0.001)
lymph_model_history = classifierObj.model_fit(epochs=10, batch_size=256)
```

Python

```
Epoch 1/10
329/329 [=====] - ETA: 0s - loss: 5.4588 - accuracy: 0.6607
Epoch 1: val_loss improved from inf to 3.79739, saving model to LymphomaVGG16.h5
329/329 [=====] - 78s 233ms/step - loss: 5.4588 - accuracy: 0.6607 - val_loss: 3.7974 - val_accuracy: 0.7236 - lr: 0.0010
Epoch 2/10
329/329 [=====] - ETA: 0s - loss: 2.0324 - accuracy: 0.8368
Epoch 2: val_loss improved from 3.79739 to 2.85107, saving model to LymphomaVGG16.h5
329/329 [=====] - 66s 201ms/step - loss: 2.0324 - accuracy: 0.8368 - val_loss: 2.8511 - val_accuracy: 0.8080 - lr: 0.0010
Epoch 3/10
329/329 [=====] - ETA: 0s - loss: 1.4824 - accuracy: 0.8795
Epoch 3: val_loss improved from 2.85107 to 2.78445, saving model to LymphomaVGG16.h5
329/329 [=====] - 65s 198ms/step - loss: 1.4824 - accuracy: 0.8795 - val_loss: 2.7845 - val_accuracy: 0.8229 - lr: 0.0010
Epoch 4/10
329/329 [=====] - ETA: 0s - loss: 1.2856 - accuracy: 0.9070
Epoch 4: val loss improved from 2.78445 to 2.64358, saving model to LymphomaVGG16.h5
```

#Oral Cancer

```
target_class = 'Oral Cancer'
target_data_path = f'/content/Multi_Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObj = ImageClassifier(no_of_classes=class_count, class_name=target_class, image_size=[224, 224])
oral_model = classifierObj.initiate_model()
```

Total : Found 10002 files belonging to 2 classes.

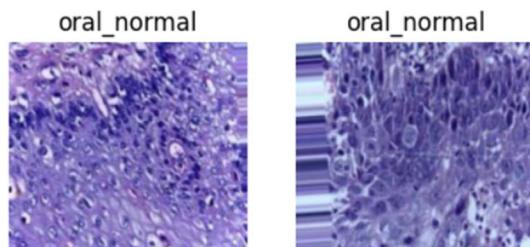
For Training : Found 7002 images belonging to 2 classes.

For Validation : Found 3000 images belonging to 2 classes.

No of Classes : 2

Classes : ['oral_normal', 'oral_scc']

Image Shape : (32, 256, 256, 3)



```
oral_model, oral_annealer, oral_model_checkpoints = classifierObj.initiate_params(lr=0.001)
oral_model_history = classifierObj.model_fit(epochs=10, batch_size=256)

Epoch 1/10
219/219 [=====] - ETA: 0s - loss: 6.6487 - accuracy: 0.7398
Epoch 1: val_loss improved from inf to 4.59375, saving model to Oral CancerVGG16.h5
219/219 [=====] - 50s 223ms/step - loss: 6.6487 - accuracy: 0.7398 - val_loss: 4.5937 - val_accuracy: 0.8057 - lr: 0.0010
Epoch 2/10
219/219 [=====] - ETA: 0s - loss: 2.4857 - accuracy: 0.8892
Epoch 2: val_loss did not improve from 4.59375
219/219 [=====] - 49s 222ms/step - loss: 2.4857 - accuracy: 0.8892 - val_loss: 5.0931 - val_accuracy: 0.8260 - lr: 0.0010
Epoch 3/10
219/219 [=====] - ETA: 0s - loss: 1.4420 - accuracy: 0.9254
Epoch 3: val_loss did not improve from 4.59375
219/219 [=====] - 40s 180ms/step - loss: 1.4420 - accuracy: 0.9254 - val_loss: 5.2399 - val_accuracy: 0.8393 - lr: 0.0010
Epoch 4/10
219/219 [=====] - ETA: 0s - loss: 1.2701 - accuracy: 0.9380
Epoch 4: val_loss did not improve from 4.59375
219/219 [=====] - 39s 180ms/step - loss: 1.2701 - accuracy: 0.9380 - val_loss: 6.0228 - val_accuracy: 0.8350 - lr: 0.0010
Epoch 5/10
219/219 [=====] - ETA: 0s - loss: 0.8965 - accuracy: 0.9530
Epoch 5: val_loss did not improve from 4.59375
219/219 [=====] - 40s 181ms/step - loss: 0.8965 - accuracy: 0.9530 - val_loss: 8.7767 - val_accuracy: 0.7863 - lr: 0.0010
Epoch 6/10
219/219 [=====] - ETA: 0s - loss: 0.7559 - accuracy: 0.9593
Epoch 6: val_loss did not improve from 4.59375
219/219 [=====] - 40s 182ms/step - loss: 0.7559 - accuracy: 0.9593 - val_loss: 7.1732 - val_accuracy: 0.8253 - lr: 0.0010
Epoch 7/10
...
Epoch 10/10
219/219 [=====] - ETA: 0s - loss: 0.7463 - accuracy: 0.9676
Epoch 10: val_loss did not improve from 4.59375
219/219 [=====] - 40s 180ms/step - loss: 0.7463 - accuracy: 0.9676 - val_loss: 9.5869 - val_accuracy: 0.8303 - lr: 0.0010
```

This code defines a class EfficientNetClassifier that encapsulates functionality for creating and training an image classifier using the EfficientNetB0 architecture in TensorFlow/Keras. It allows for easy customization of the number of classes, input image size, and data generators for training and validation.

```
from tensorflow.keras.applications import EfficientNetB0
from tensorflow.keras.layers import Flatten, Dense
from tensorflow.keras.models import Model
from tensorflow.keras.callbacks import ReduceLROnPlateau, ModelCheckpoint
import tensorflow as tf

class EfficientNetClassifier:
    def __init__(self, no_of_classes, image_size, class_name, train_generator,
validation_generator):
        self.no_of_classes = no_of_classes
        self.image_size = image_size
        self.class_name = class_name
        self.train_generator = train_generator
        self.validation_generator = validation_generator
        self.model = None
        self.annealer = None
        self.checkpoint = None

    def initiate_model(self):
        model_input = EfficientNetB0(
            input_shape=self.image_size + [3],
            include_top=False,
            weights="imagenet"
        )

        for layer in model_input.layers:
            layer.trainable = False

        x = Flatten()(model_input.output)
        prediction = Dense(self.no_of_classes, activation='relu')(x)
```

```

# Create the model object
self.model = Model(inputs=model_input.input, outputs=prediction)
return self.model

def model_summary(self):
    if self.model is not None:
        self.model.summary()
    else:
        print("Model has not been initialized yet.")

def initiate_params(self, lr):
    if self.model is None:
        print("Model has not been initialized yet.")
    return

opt = tf.keras.optimizers.Adam(learning_rate=lr)
self.model.compile(optimizer=opt, loss='categorical_crossentropy', metrics=['accuracy'])

self.annealer = ReduceLROnPlateau(monitor='val_accuracy', factor=0.5, patience=5,
verbose=1, min_lr=1e-3)
self.checkpoint = ModelCheckpoint(self.class_name + 'EfficientNetB0.h5', verbose=1,
save_best_only=True)

return self.model, self.annealer, self.checkpoint

def model_fit(self, epochs=20, batch_size=256):
    if self.model is None:
        print("Model has not been initialized yet.")
    return

history = self.model.fit(
    self.train_generator,
    validation_data=self.validation_generator,

```

```

    epochs=epochs,
    batch_size=batch_size,
    callbacks=[self.annealer, self.checkpoint],
    steps_per_epoch=len(self.train_generator),
    validation_steps=len(self.validation_generator)
)
return history

def eval_model(self):
    if self.model is None:
        print("Model has not been initialized yet.")
        return

    # Evaluate the model
    evl = self.model.evaluate(self.validation_generator)
    acc = evl[1] * 100
    msg = f'Accuracy on the Test Set = {acc:5.2f} %'
    print(msg)

def save_model(self):
    if self.model is None:
        print("Model has not been initialized yet.")
        return

    file_path = self.class_name + " - EfficientNetB0.h5"
    self.model.save(file_path)
    print(f"Model saved to {file_path}!")

target_class = 'Cervical Cancer'
target_data_path = f'/content/Multi Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

```

```

classifierObjB0 = EfficientNetClassifier(no_of_classes=class_count, class_name=target_class,
image_size=[224, 224], train_generator=train_gen, validation_generator=valid_gen)
cervical_modelB0 = classifierObjB0.initiate_model()
cervical_modelB0, cervical_annealerB0, cervical_model_checkpointsB0 =
classifierObjB0.initiate_params(lr=1)
cervical_model_history_B0 = classifierObjB0.model_fit(epochs=10, batch_size=256)
brain_modelB0, brain_annealerB0, brain_model_checkpointsB0 =
classifierObjB0.initiate_params(lr=1)
brain_model_history_B0 = classifierObjB0.model_fit(epochs=10, batch_size=256)
#Cervical Cancer
target_class = 'Cervical Cancer'
target_data_path = f'/content/Multi Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObjB0 = EfficientNetClassifier(no_of_classes=class_count, class_name=target_class,
image_size=[224, 224], train_generator=train_gen, validation_generator=valid_gen)
kidney_modelB0 = classifierObjB0.initiate_model()
kidney_modelB0, kidney_annealerB0, kidney_model_checkpointsB0 =
classifierObjB0.initiate_params(lr=1)
kidney_model_history_B0 = classifierObjB0.model_fit(epochs=10, batch_size=256)

#Breast Cancer
target_class = 'Cervical Cancer'
target_data_path = f'/content/Multi Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObjB0 = EfficientNetClassifier(no_of_classes=class_count, class_name=target_class,
image_size=[224, 224], train_generator=train_gen, validation_generator=valid_gen)
breast_modelB0 = classifierObjB0.initiate_model()

```

```

target_class = 'Cervical Cancer'
target_data_path = f'/content/Multi Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObjB0 = EfficientNetClassifier(no_of_classes=class_count, class_name=target_class,
image_size=[224, 224], train_generator=train_gen, validation_generator=valid_gen)
lung_modelB0 = classifierObjB0.initiate_model()
lung_modelB0, lung_annealerB0, lung_model_checkpointsB0 =
classifierObjB0.initiate_params(lr=1)
lung_model_history_B0 = classifierObjB0.model_fit(epochs=10, batch_size=256)
target_class = 'Cervical Cancer'
target_data_path = f'/content/Multi Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

classifierObjB0 = EfficientNetClassifier(no_of_classes=class_count, class_name=target_class,
image_size=[224, 224], train_generator=train_gen, validation_generator=valid_gen)
lymph_modelB0 = classifierObjB0.initiate_model()

lymph_modelB0, lymph_annealerB0, lymph_model_checkpointsB0 =
classifierObjB0.initiate_params(lr=1)
lymph_model_history_B0 = classifierObjB0.model_fit(epochs=10, batch_size=256)

target_class = 'Cervical Cancer'
target_data_path = f'/content/Multi Cancer/{target_class}'

dataProcessor = ImageDataProcessor(target_data_path)

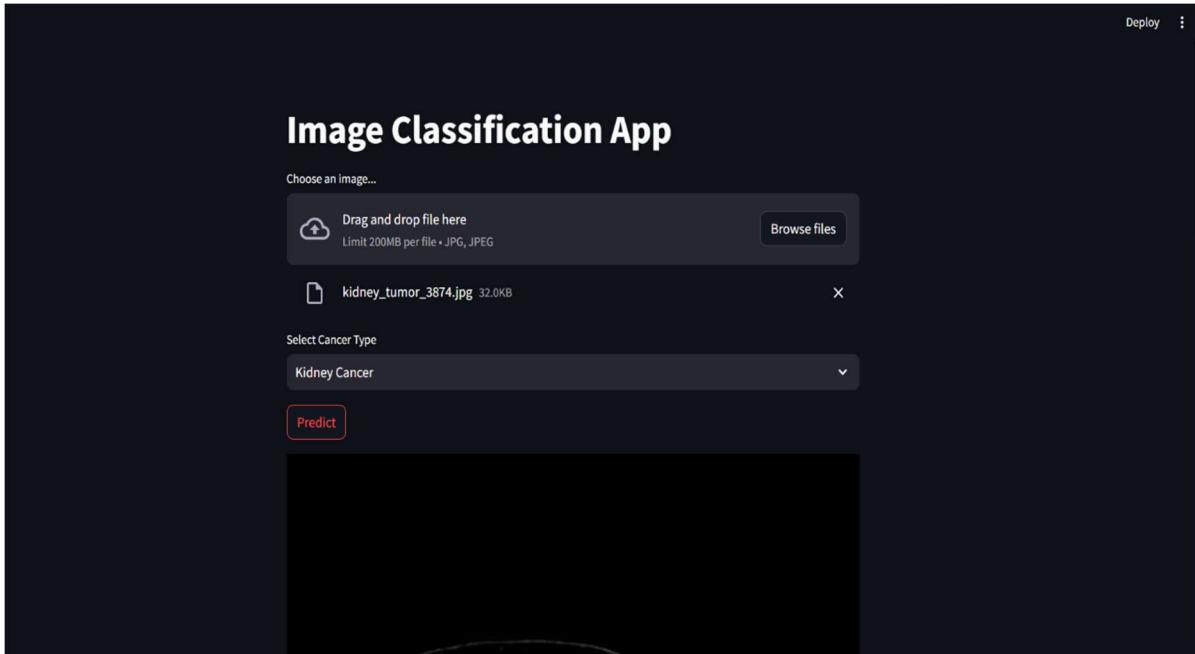
```

```
classes, class_count, train_gen, valid_gen = dataProcessor.initiate_generator()

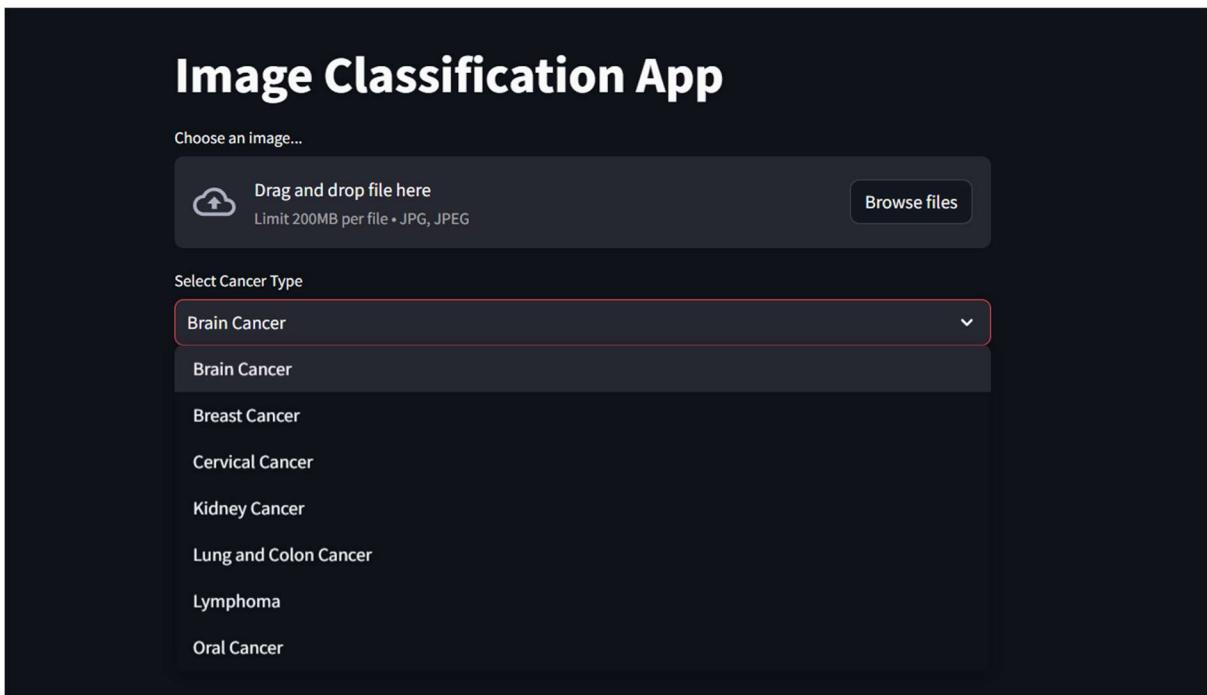
classifierObjB0 = EfficientNetClassifier(no_of_classes=class_count, class_name=target_class,
image_size=[224, 224], train_generator=train_gen, validation_generator=valid_gen)
oral_modelB0 = classifierObjB0.initiate_model()
oral_modelB0, oral_annealerB0, oral_model_checkpointsB0 =
classifierObjB0.initiate_params(lr=1)
oral_model_history_B0 = classifierObjB0.model_fit(epochs=10, batch_size=256)
```

APPENDIX 2

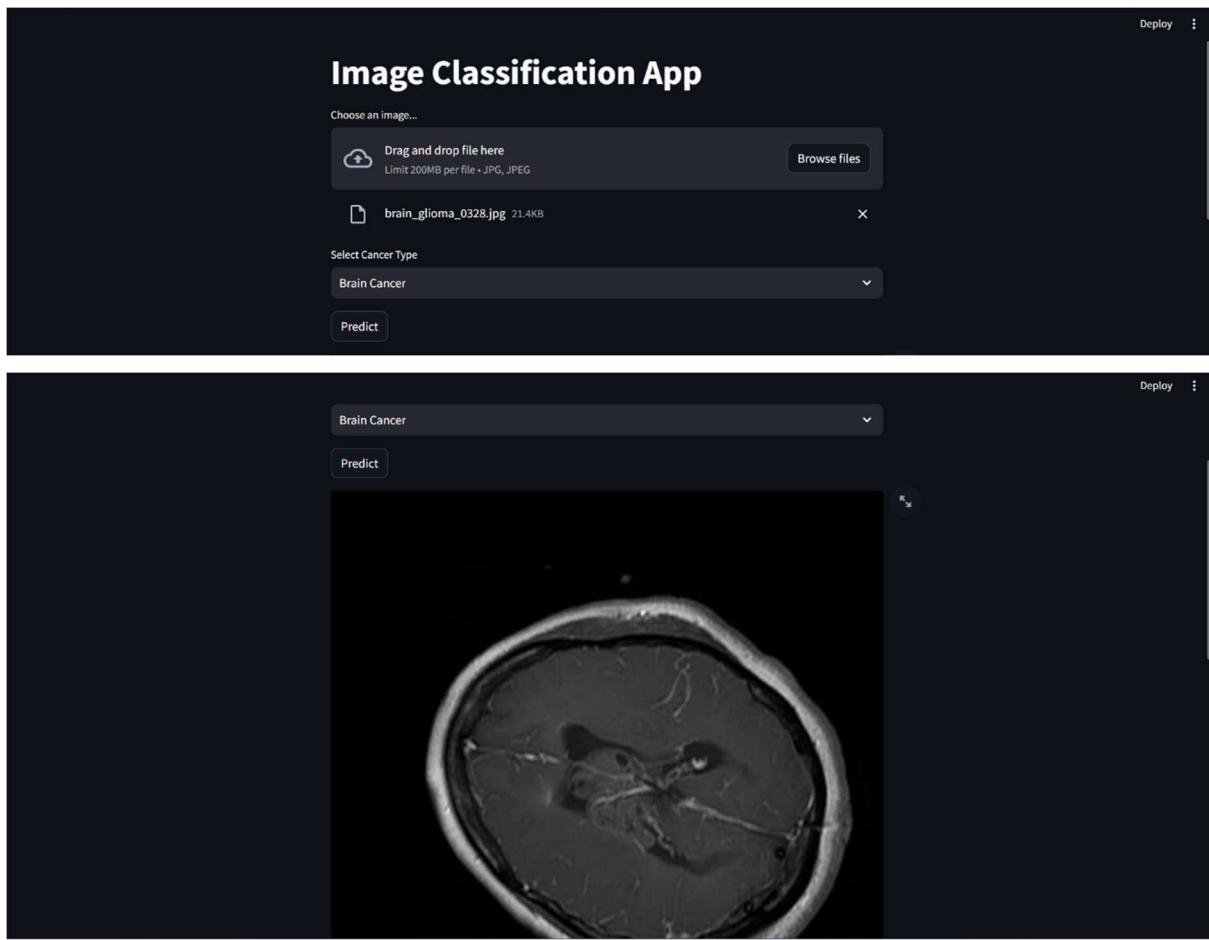
Output:



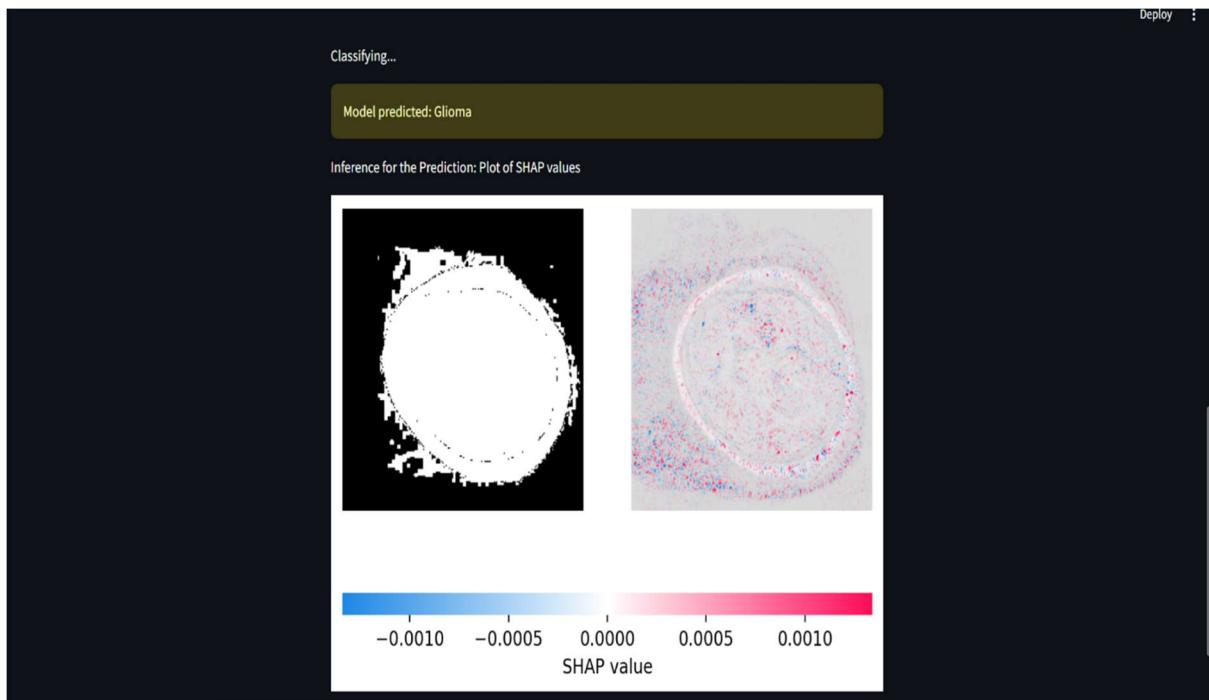
Home page



Selecting cancer type and uploading image



Selecting the type of cancer and uploading image



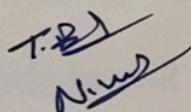
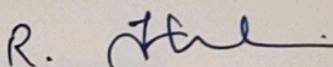
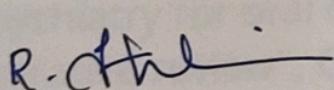
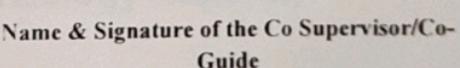
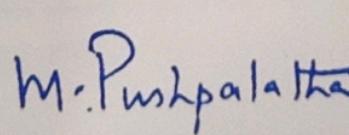
Result Final Model Predicted

PLAGIARISM REPORT

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY		
(Deemed to be University u/ s 3 of UGC Act, 1956)		
Office of Controller of Examinations		
REPORT FOR PLAGIARISM CHECK ON THE DISSERTATION/PROJECT REPORTS FOR UG/PG PROGRAMMES		
(To be attached in the dissertation/ project report)		
1	Name of the Candidate (IN BLOCK LETTERS)	TALLURU VENKATA BHUVANESH
2	Address of the Candidate	Nakshatra Apartment Room no-7, Chamundeshwari Nagar, Thailavaram TN – 603 203.
3	Registration Number	RA2011003011298
4	Date of Birth	25-02-2003
5	Department	Computer Science and Engineering
6	Faculty	Engineering and Technology, School of Computing
7	Title of the Dissertation/Project	BLOCK- A Multi Cancer Disease Detection
8	Whether the above project /dissertation is done by	<p style="margin-left: 20px;">Individual or group : (Strike whichever is not applicable)</p> <p style="margin-left: 20px;">a) If the project/ dissertation is done in group, then how many students together completed the project : 2</p> <p style="margin-left: 20px;">b) Mention the Name & Register number of other candidates : NANDIMANDALAM VIVEK [RA2011003011302].</p>
9	Name and address of the Supervisor / Guide	DR. R. THENMOZHI Associate Professor Department Of Computing Technologies Kattankulathur – 603 203 Mail ID: 8838749603 Mobile Number: thenmozr@srmist.edu.in
10	Name and address of Co-Supervisor / Co- Guide (if any)	NA

11	Software Used	Turnitin		
12	Date of Verification	25-04-2024		
13	Plagiarism Details: (to attach the final report from the software)			
Chapter	Title of the Chapter	Percentage of similarity index (including self citation)	Percentage of similarity index (Excluding self-citation)	% of plagiarism after excluding Quotes, Bibliography, etc.,
1	INTRODUCTION	1%	1%	1%
2	LITERATURE SURVEY	2%	2%	2%
3	SYSTEM ARCHITECTURE AND DESIGN	1%	1%	1%
4	METHODOLOGY	3%	3%	3%
5	RESULT AND DISCUSSION	1%	1%	1%
6	CONCLUSION AND FUTURE ENHANCEMENT	0%	0%	0%
Appendices		1%	1%	1%

I / We declare that the above information have been verified and found true to the best of my / our knowledge.

 Signature of the Candidate	 Name & Signature of the Staff(Who uses the plagiarism check software)
 Name & Signature of the Supervisor/ Guide	 Name & Signature of the Co Supervisor/Co-Guide
 Name & Signature of the HOD	



9
%

SIMILARITY INDEX

5
%

INTERNET SOURCES

5
%

PUBLICATIONS

2
%

STUDENT PAPERS

PRIMARY SOURCES

1

Malliga Subramanian, Jaehyuk Cho,
Veerappampalayam Easwaramoorthy
Sathishkumar, Obuli Sai Naren. "Multiple
Types of Cancer Classification Using CT/MRI
Images Based on Learning Without Forgetting
Powered Deep Learning Models", IEEE
Access, 2023

1
%

Publication

2

assets.researchsquare.com

1
%

Internet Source

3

fastercapital.com

1
%

Internet Source

4

Abdulmalik Fareeq, Sirwan Khalid Ahmed,
Safin Hussein, Karzan Qurbani. "Artificial
intelligence-assisted nursing interventions in
psychiatry for oral cancer patients: A concise
narrative review", Oral Oncology Reports,
2024

<1
%

Publication

5

www.science.gov

<1
%

Internet Source

-
- 7 Somit Jain, Dharmik Naicker, Ritu Raj, Vedanshu Patel, Yuh-Chung Hu, Kathiravan Srinivasan, Chun-Ping Jen. "Computational Intelligence in Cancer Diagnostics: A Contemporary Review of Smart Phone Apps, Current Problems, and Future Research Potentials", Diagnostics, 2023
Publication <1 %
-
- 8 www.hindawi.com <1 %
Internet Source
-
- 9 Submitted to Asia Pacific University College of Technology and Innovation (UCTI) <1 %
Student Paper
-
- 10 Submitted to Liverpool John Moores University <1 %
Student Paper
-
- 11 www.ijisae.org <1 %
Internet Source
-
- 12 Submitted to University of Wales Institute, Cardiff <1 %
Student Paper
-
- 13 Viswambari Devi Ramaswamy, Michael Keidar. "Personalized Plasma Medicine for Cancer: Transforming Treatment Strategies with <1 %

Publication

14	Submitted to universititeknologimara Student Paper	<1 %
15	vdoc.pub Internet Source	<1 %
16	www.nursa.org Internet Source	<1 %
17	Yang Yu, Bhavya Jain, Gautam Anand, Mahdi Heidarian, Andrew Lowe, Anubha Kalra. "Technologies for non-invasive physiological sensing: Status, challenges, and future horizons", Biosensors and Bioelectronics: X, 2024 Publication	<1 %
18	Submitted to George Bush High School Student Paper	<1 %
19	documents.mx Internet Source	<1 %
20	marketingeye27.blogspot.com Internet Source	<1 %
21	www.digitaljournal.com Internet Source	<1 %
22	Submitted to Concordia University Student Paper	<1 %

24	Submitted to University of Glamorgan Student Paper	<1 %
25	filmdaily.co Internet Source	<1 %
26	ouci.dntb.gov.ua Internet Source	<1 %
27	www.coherentmarketinsights.com Internet Source	<1 %
28	www.ijrcog.org Internet Source	<1 %
29	Submitted to Midlands State University Student Paper	<1 %
30	Submitted to University of Northampton Student Paper	<1 %
31	Usman Yunus, Javeria Amin, Muhammad Sharif, Mussarat Yasmin, Seifedine Kadry, Sujatha Krishnamoorthy. "Recognition of Knee Osteoarthritis (KOA) Using YOLOv2 and Classification Based on Convolutional Neural Network", Life, 2022 Publication	<1 %
32	www.researchgate.net Internet Source	<1 %

-
- 34 Shabnam Khatami, Lei Xuan, Rolando Roman, Song Zhang, Charles McConnel, Ethan A. Halm, Samir Gupta. "Modestly Increased Use of Colonoscopy When Copayments Are Waived", *Clinical Gastroenterology and Hepatology*, 2012
Publication <1 %
-
- 35 Submitted to cnsc Student Paper <1 %
-
- 36 export.arxiv.org Internet Source <1 %
-
- 37 www.eurekalert.org Internet Source <1 %
-
- 38 Humaira Shafiq, Ghulam Gilanie, Muhammad Sajid, Muhammad Ahsan. "Dental radiology: a convolutional neural network-based approach to detect dental disorders from dental images in a real-time environment", *Multimedia Systems*, 2023
Publication <1 %
-
- 39 Kaijiong Zhang, Bo Ye, Lichun Wu, Sujiao Ni, Yang Li, Qifeng Wang, Peng Zhang, Dongsheng Wang. "Machine learning-based prediction of survival prognosis in esophageal <1 %

Publication

40	saucis.sakarya.edu.tr	<1 %
41	www.ajmc.com	<1 %
42	www.researchsquare.com	<1 %
43	Behavioural Oncology, 2014. Publication	<1 %
44	Marta Bertolaso. "Philosophy of Cancer", Springer Science and Business Media LLC, 2016 Publication	<1 %
45	Zhengqian Xu, Peiying Zhang, Chengcheng Li, Hailong Zhu, Guanjun Xu, Chenhua Sun. "A Collaborative Inference Algorithm in Low- Earth-Orbit Satellite Network for Unmanned Aerial Vehicle", Drones, 2023 Publication	<1 %
46	m.giikorea.co.kr	<1 %
47	worldwidescience.org	<1 %

www.trademarkelite.com

49

"Neural Information Processing", Springer
Science and Business Media LLC, 2017
Publication

<1 %

Exclude quotes Off
Exclude bibliography Off

Exclude matches Off

PAPER PUBLICATION PROOF

Our Research paper titled “Treatment Recommendation For Medical Conditions” has been accepted for presentation at **the international conference on Recent Advances in Computer Science and information Technology**. This prestigious conference is organised by **TECHOWN**, a renowned institution committed to promoting advancements in technology and innovation. We have received the Acknowledgement that our paper has been successfully Accepted. The conference is held on 21st April 2024 in Chennai.

4/23/24, 6:55 PM SRM Institute of Science and Technology Mail - Subject: Acceptance of Research paper for International Conference on Rece...



TALLURU VENKATA BHUVANESH (RA2011003011298) <tb5348@srmist.edu.in>

Subject: Acceptance of Research paper for International Conference on Recent Advances in Computer Science and Information Technology

info@techown.in <info@techown.in> Wed, Mar 20, 2024 at 12:15 AM
To: tb5348@srmist.edu.in, Nn3007@srmist.edu.in, Vd3023 <vd3023@srmist.edu.in>

Dear Venkata Bhuvanesh Talluru, Vivek Nandimandalam, Thenmozhi R, Vetriselvi D,

We are thrilled to inform you that your research paper titled “Treatment Recommendation For Medical Conditions” has been accepted for presentation at the International Conference on Recent Advances in Computer Science and Information Technology, scheduled to be held on 21st April 2024 in Chennai. This prestigious conference is organized by TECHOWN, a renowned institution committed to promoting advancements in technology and innovation.

Your contribution to this conference will be instrumental in enriching the discourse surrounding the latest developments in computer science and information technology. Your research paper was chosen based on its quality, significance, and potential impact on the field.

The conference will commence with registration starting at 8:00 AM on the day of the event. The program will include various sessions, encompassing keynote addresses, paper presentations, panel discussions, and networking opportunities. A detailed agenda will be provided closer to the conference date.

As a presenter, you will have the opportunity to share your insights and findings with an audience of fellow researchers, academicians, and industry professionals. Your participation will foster collaboration and facilitate the exchange of ideas within the global research community.

All registered attendees will receive certificates of participation, and outstanding presentations will be recognized during the conference proceedings. We encourage you to prepare for your presentation and actively engage with other participants to maximize the benefits of this scholarly gathering.

Should you have any inquiries or require further information, please feel free to contact us at info@techown.in. We are here to assist you in any way possible.

Once again, congratulations on the acceptance of your research paper. We eagerly anticipate your contribution to the success of the International Conference on Recent Advances in Computer Science and Information Technology.

Warm regards,

Ashok Kumar. P
Conference Coordinator: TECHOWN.