AI Driven Ovarian Cancer Early Detection by Subtype Categorization and Aberrant Instances Identification

by

Ishtiaq Ahmed Shuvo
20301429
Masrurah Madeeha
20201113
B M Nayeem Samad
20301251

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

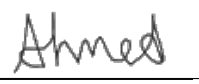
Department of Computer Science and Engineering
Brac University
June 2025

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

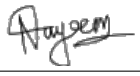


Ishtiaq Ahmed Shuvo

ID: 20301429

Masrurah Madeeha

ID: 20201113



B M Nayeem Samad

ID: 20301251

Approval

The thesis titled "AI Driven Ovarian Cancer Early Detection by Subtype Categorization and Aberrant Instances Identification" submitted by

1. Ishtiaq Ahmed Shuvo (20301429)
2. Masrurah Madeeha (20201113)
3. B M Nayeem Samad (20301251)

Of Spring, 2025 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on June 20, 2025.

Examining Committee:

Supervisor:
(Member)

Jawaril Munshad Abedin

Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Dr. Sadia Hamid Kazi

Associate Professor and Chairperson
Department of Computer Science and Engineering
Brac University

Ethics Statement

The proposed paper is an amazing endeavor. Furthermore, the members hereby and genuinely affirm that this was done based on the findings of our thorough investigation. All resources used are properly noted and credited in the report. This research work, in whole or in part, has never been submitted to another university or institution for degree conferring or for any other purpose.

Abstract

Among cancers of the female reproductive system, ovarian cancer stands as one of the most deadly types because of its ability to vary between patients and its late discovery by medical professionals. The correct identification of ovarian cancer subtypes serves both therapeutic purposes and outcome enhancement needs in medical treatment. This study presents an automated, robust classification system based on advanced machine learning architectures to overcome the challenges of subtype classification. Research on ovarian cancer subtypes still has not progressed until this work since class imbalance and uncommon subtypes were often neglected. Previously, many studies functioned under basic models and single-source data without which accuracy and generalization are limited. Specifically, performance evaluation overemphasized only accuracy while overrunning important metrics like precision and AUC. Data augmentation was underused, and early detection was often neglected. This study fills the voids to enhance diagnostic capabilities in ovarian cancer. Several deep learning architectures are considered for improving the performance of classification in this work, including pre-trained architectures like VGG16 and ResNet50, a hybrid model by taking the strengths of both, CNN, and Vision Transformers. Our hybrid model performed the best with an accuracy of 92.8%, F1-score of 0.927, recall of 92.8%, and precision of 93.5%, which showed that the integration of complementary feature extraction capabilities was effective. Other advanced data preprocessing techniques, like resizing, normalization, and augmentation, are also employed in this study to improve model generalization and handle class imbalance. This work provides a very efficient yet scalable classification framework in the field of medical imaging and diagnostics of cancer. The results have pointed out the importance of hybrid architectures and pre-trained models toward superior performance and this system has great potential to be integrated into a clinical workflow; it provides a useful tool to support pathologists and oncologists in the diagnosis of ovarian cancer.

Keywords: Ovarian cancer, subtype classification, machine learning, hybrid model, VGG16, ResNet50, Vision Transformer, medical imaging, cancer diagnosis, histopathological images.

Dedication

This work is dedicated to all who inspired and supported us on the journey of research and development. We must express our deep appreciation to our families, whose unwavering support and encouragement have remained our strongest strength and motivational force. We are deeply grateful for the guidance, wisdom, and belief of our mentors and teachers regarding our potential throughout our academic and professional growth. Finally, this research is dedicated to the medical and scientific community, whose commitment to the betterment of healthcare outcomes and advancement of knowledge continues to inspire innovation and progress in the fight against ovarian cancer.

Acknowledgement

To begin, we would want to offer all appreciation to the Almighty, because of whom our thesis was finished without any significant interruptions. Then we want to acknowledge our supervisor without her valuable guidance and feedback we would not be able to come this far.

Table of Contents

Declaration i

Approval ii

Ethics Statement iii

Abstract iv

Dedication v

Acknowledgment vi

Table of Contents vii

List of Figures ix

List of Tables x

1 Introduction 1
- 1.1 Problem Statement . . . 3
- 1.2 Research Questions . . . 4
- 1.3 Objectives . . . 5
- 1.4 Scope of study . . . 6

2 Literature Review 8
- 2.1 Detailed Literature Review . . . 8
- 2.2 Research Gaps . . . 13

3 The Dataset 15
- 3.1 Overview of the Dataset . . . 15
- 3.2 Methodology . . . 17
- 3.3 Data Preprocessing . . . 19
  - 3.3.1 Data Loading and Organization . . . 19
  - 3.3.2 Data Splitting . . . 19
  - 3.3.3 Resizing . . . 20
  - 3.3.4 Rescaling . . . 20
  - 3.3.5 Data Augmentation . . . 20
- 3.4 Optimization and Fine-Tuning . . . 21
  - 3.4.1 Layer Freezing and Fine-Tuning . . . 21
  - 3.4.2 Batch Size Adjustments . . . 21

3.4.3 Learning Rate Adjustments . . . 21
3.4.4 Additional Regularization Techniques . . . 22
3.4.5 Evaluation and Fine-Tuning on Validation Set . . . 22
3.5 Data Splitting for Model Training . . . 22
3.5.1 Importance of Data Splitting . . . 22
3.5.2 Stratified Splitting . . . 22
3.5.3 Randomization and Reproducibility . . . 23
3.5.4 Handling Data Imbalance . . . 23
3.5.5 Validation Strategy . . . 23
3.5.6 Testing Set Considerations . . . 23
3.5.7 Challenges in Data Splitting . . . 23
3.6 Ethical Considerations . . . 24

4 Methodology **25**
4.1 Research Framework . . . 25
4.2 Description of the Model . . . 26
4.2.1 Convolutional Neural Network (CNN) . . . 26
4.2.2 Vision Transformers(ViT) . . . 27
4.2.3 VGG-16 . . . 28
4.2.4 Residual Network(ResNet50) . . . 29
4.2.5 Hybrid Model (VGG16 + ResNet50) . . . 30
4.3 Model Training and Evaluation . . . 31

5 Result **32**
5.1 Evaluation Metrics . . . 32
5.2 Performance Analysis of Different Architectures . . . 35
5.2.1 Discussion of Results . . . 35
5.2.2 Key Insights . . . 37

6 Conclusion **38**
6.1 Conclusion of the study . . . 38
6.2 Significance of the Research . . . 39
6.3 Future Works . . . 39

Bibliography **41**

List of Figures

3.1 Samples from dataset 16
3.2 Workflow of the Methodology 18
3.3 Gantt Chart 18
3.4 Workflow of Data Preprocessing 19

4.1 CNN Architecture 26
4.2 ViT Architecture 28
4.3 VGG16 Architecture [1] 28
4.4 ResNet50 Architecture [2] 29
4.5 Hybrid Model Architecture [3] 31

5.1 CNN Confusion Matrix 32
5.2 VGG16 Confusion Matrix 33
5.3 ResNet Confusion Matrix 33
5.4 Hybrid Model Confusion Matrix 34
5.5 Predicted Classes Realtime 34
5.6 Accuracy and Loss graph of CNN Model 35
5.7 Accuracy and Loss graph of VGG16 Model 36
5.8 Accuracy and Loss graph of ResNet50 Model 36
5.9 Accuracy and Loss graph of Hybrid Model 37

List of Tables

3.1 Ovarian Cancer Subtype Distribution in the UBC-OCEAN Dataset . 15

4.1 Comparison of Model Parameters, Epochs, and Training Times 30

5.1 Performance Comparison of Different Architectures 35

Chapter 1

Introduction

Ovarian cancer remains one of the most lethal cancers that has an impact on the female reproductive system, posing a significant global health challenge [4]. It has mostly been diagnosed when it reaches an advanced stage, when treatment outcomes are not very favorable. In 2020 alone, the World Health Organization(WHO) reported approximately 313,959 new ovarian cancer cases and 207,252 deaths worldwide [4]. Ovarian cancer is not a single disease but it is a clinically and molecularly heterogeneous malignancy. In addition, Ovarian cancer is a group of cancers classified into five main subtypes: High-grade Serous Carcinoma (HGSC), Clear Cell Carcinoma (CC), Endometrioid Carcinoma (EC), Low-grade Serous Carcinoma (LGSC), and Mucinous Carcinoma (MC). Each subtype has its own unique molecular profiles and clinical behavior, along with cellular morphology [4]. Hence, accurate classification is the only key to offering a personalized treatment plan for enhancing the outcome. Identifying subtypes is clinically critical as their symptoms vary based on prognoses and therapeutic responses. For instance, Clear Cell Carcinoma(CC) tends to have a higher recurrence rate and poorer prognosis than the common high-grade serous carcinoma subtypes [5]. Such behavioral differentiation between these subtypes makes it crucial to provide precise classifications for personalized treatment decisions. The traditional approach to identifying an ovarian tumor's type, done by expert pathologists, is regarded as the best method, though it also has its difficulties and limitations. Usually, expert pathologists examine tissue to determine the subtype of an ovarian tumor, where they rely on subjective visual examination of microscopy slides, but this might deliver different results based on their mood and observational mistakes. Identifying subtypes is a challenging task using microscopy slides, as they often show similar patterns in their histopathology images. Also, it becomes nearly impossible to differentiate them at an early stage [4]. A delicate morphological distinction between subtypes can be challenging to understand even for an experienced pathologist, which may lead to potential disagreement in diagnosis [4]. Therefore, manual slide review is hard and time-consuming, which causes a delay in the diagnosis procedure. These challenges create a necessity for assistive diagnostic tools that can provide objective and reproducible subtype classification. With a reliable automated system, doctors can identify the subtype of ovarian cancer and start the best therapy for each patient quickly. Deep learning and artificial intelligence can provide a potential way to improve the classification of ovarian cancer in histopathology. Particularly, convolutional neural networks (CNNs) have stood out by often detecting patterns in medical images with more

precision than humans can achieve [4][6]. Deep learning can enhance image analysis by spotting small differences in cellular patterns and parts of tissue arising from diseases [4]. CNN architecture has been recently used in ovarian cancer pathology to great advantage. Scientists have tried several models, including VGG16, ResNet50, DenseNet, Xception, MobileNetV2, and EfficientNet, for ovarian tumor detection or classification, and each has proven high accuracy [7][4]. To illustrate, using an enhanced ResNet50 achieved 97.5% accuracy for classifying ovarian tumors, which was better than several modern models [7]. The same can be said for EfficientNet-based classifiers, which have achieved nearly perfect accuracy($\sim$100%) using limited histological images from the ovarian biopsies [4]. These successes highlight the potential of deep learning to diagnose ovarian cancer. However, many prior works have addressed only binary classification (distinguishing malignant vs benign ovarian tumors) or focused on a single subtype rather than analyzing the multi-class subtype classification [4]. Due to the shortage of large, structured datasets and detailed model comparisons, it remains difficult to improve solutions suitable for real-world care [4]. New studies are needed to classify multi-class ovarian cancers and to compare different machine learning models on this task unbiasedly. The approach we present in this study combines two types of CNNs to classify ovarian cancer into different subtypes. In this research, we focus on identifying the five major subtypes of epithelial ovarian carcinoma (HGSC, CC, EC, LGSC, MC) in a multi-class classification framework[4]. In this research, we proposed a design that merges the feature extraction parts of VGG16 and ResNet50, utilizing the positives of both networks. Before that, we implemented Base CNN and ViT and also tried customizing layers to see the result using the same dataset. Also tested ResNet50 and VGG16 individually. VGG16 is popular because its structure is detailed and similar all the way through, allowing it to notice small textures, while ResNet50 uses skip connections, which help the network use more complex features without dealing with vanishing gradients. Combining VGG16 and ResNet50 allows the hybrid model to spot more varieties of patterns in histopathology than just one could. When using such strategies, accuracy increases, and the risk of overfitting in medical image classification is reduced [8]. To the best of our knowledge, this study is one of the first to use a hybrid CNN combining VGG16 and ResNet50 for multi-class identification of ovarian carcinoma types. All of our research and analysis were done on computerized H&E pictures, and we measured the model's performance alongside several others. The hybrid approach was better at identifying the subtypes of ovarian cancer. The accuracy was achieved with 92.8.% accuracy and an F1-score of 0.927, far surpassing the accuracy of ResNet50 and VGG16 (both achieved $\sim$84.% accuracy in our tests individually). Integrating features from various CNNs led to an improvement of 9 percentage points in the accuracy. This research evaluated our results against recently published advanced neural network models. DenseNet, MobileNetV2, Xception, and EfficientNet, which are leading deep CNNs, reached classification accuracy levels of 90–99.% for diagnosing ovarian histopathology cases according to published literature [4][7][9]. In contrast to many prior studies that rely on extensive data preprocessing, selective sampling, or dataset reduction techniques, our approach demonstrates that high classification reliability can be achieved directly on the raw, unfiltered UBC-OCEAN dataset. We intentionally did not divide or reduce the data, so it still reflects everything seen in these hospitals. By doing so, we aimed to maintain data integrity while developing a model that could

generalize effectively across all five major ovarian carcinoma subtypes. Despite these constraints, our proposed hybrid model (VGG16 + ResNet50) achieved the highest accuracy and F1-score among all tested architectures, including Vision Transformers and standard CNN baselines. In addition, we worked to find a balance between our algorithm's accuracy and its computational cost, making the training more efficient. To the best of our knowledge, no existing research has attained comparable predictive consistency across all subtypes using the raw Kaggle-provided dataset without preprocessing shortcuts or clinical feature enhancements. The hybrid model works equally well or better than these architectures for subtype classification, while handling more challenging tasks of five-class differentiation. Unlike certain previous studies that have relied on smaller or limited data, we applied our model to a wide range of ovarian tumor images, including all the main subtypes, which makes our findings more useful for clinical relevance. These results suggest that the hybrid model might be a reliable choice for making pathological diagnoses. What matters is the promise of this study to improve patient care. A system using AI to group ovarian cancer types could improve the way pathologists handle their duties.. By providing rapid and consistent second opinions on tumor subtype, this study can help mitigate problems caused by mistakes and variations in observation among pathologists [4]. The model can mark the cases which require additional attention or assist in making the diagnosis more homogenous across various hospitals, resulting in similar grouping of patients. Incorporation of the technology would facilitate clinical practice to aid in better treatment planning. Take an instance; it will make sure that a patient with clear cell carcinoma is properly identified and considered as an option for different treatment approaches. What is more, applying this method in the initial stages enables the integration of imaging-based AI with other types of data to aid in making superior decisions. Recent studies show that the diagnosis of cancer can be substantially improved by combining multi-dimensional data into cancer diagnostic models [10]. Later, our histopathology classifier could process genomic mutations, transcription profiles, and patient data to help build a complete predictive model. Putting all the data together usually does a better job than using just a single approach in forecasting cancer outcomes and categorizing them into subtypes [10]. By bringing together histopathological images and molecular biomarkers, we hope to improve the way we define subtypes and find new markers for treatment in ovarian cancer, moving us closer to precision medicine in this disease [10].

1.1 Problem Statement

Ovarian cancer is one of the most dangerous types among females, mainly because of its usually late-stage diagnosis. The implications of such lateness in detection dramatically affect the outcomes and survival; therefore, the importance of diagnosis to be as early and specific as possible cannot be underestimated. Each of these five major ovarian carcinomas- high-grade serous carcinoma, clear-cell ovarian carcinoma, endometrioid, low-grade serous carcinoma, and mucinous carcinoma- presents specific biological and clinical features, and each of these requires appropriate treatment strategies. This ultimately leads to poor patient outcomes due to suboptimal treatment plans because of misclassification or failure to identify the correct subtype. There is, therefore, an urgent need for a reliable and efficient method of subtype classification that will help the clinician provide precise and personalized care.

Traditional methods of ovarian cancer subtype classification rely heavily on histopathological examination by expert pathologists. Although effective, these techniques are generally subjective, time-consuming, and susceptible to inter-observer variability. Most importantly, all these techniques rely on human expertise, which again restricts their overall scalability, especially in areas where there is a shortage of trained pathologists. Inaccurate diagnosis not only postpones the treatment process but also contributes to the high mortality rate associated with ovarian cancer. This represents a big gap in the current clinical practice that has to be met if better quality care among ovarian cancer patients is to be observed.
Besides, heterogeneity of ovarian cancer subtypes presents unique challenges for classification. Each subtype has overlapping visual patterns in histopathology images, which cannot be correctly discriminated using traditional approaches. Moreover, the variability in tumor morphology even within the same subtype further complicates the classification process. These challenges underline the need for advanced tools able to catch subtle differences and provide consistent results even in cases that are complex and ambiguous.
Though machine learning and deep learning are found to be very successful in medical image analysis, their applications to ovarian cancer subtype classification remain unexploited. Most of the works done in this domain are binary classification tasks, namely, distinguishing malignant from benign tumors, rather than multiclass subtypes. Apart from these, the availability of non-standard datasets and non-exhaustive comparisons of models prevents current research from finding generalized and clinically relevant outcomes. Unless this deficit is overcome, one cannot really explore the full capabilities of machine learning for ovarian cancer diagnostics. In such a backdrop, the proposed research seeks to delve into advanced techniques of machine learning applied to subtype classification of ovarian cancers. In this work, robust automated classification is developed by leveraging custom neural network architectures together with pretrained models and state-of-the-art approaches such as Vision Transformers. The goal of this work has remained not only to further improve diagnostic accuracy but also to scale, so that it is clinically useful and, hopefully, further improves treatment outcomes and reduces mortality among ovarian cancer patients.

1.2 Research Questions

1. How does the proposed hybrid model combining VGG16 and ResNet50 outperform other deep learning architectures in classifying ovarian cancer subtypes?

2. What impact do preprocessing techniques, such as image resizing, normalization, and augmentation have on the classification performance of histopathological images?

3. What are the comparative strengths and limitations of the deep learning models and methods evaluated in this study for ovarian cancer subtype classification?

1.3 Objectives

The primary objectives of this research on ovarian cancer subtype classification using machine learning techniques are as follows:

1. **To create high-reliability classification model of the raw UBC-OCEAN dataset:** The current study will help develop a strong and transferable classification scheme that will be able to identify the five key histopathological subtypes of ovarian cancer, including high-grade serous carcinoma (HGSC), clear cell carcinoma (CC), endometrioid carcinoma (EC), low-grade serous carcinoma (LGSC), and mucinous carcinoma (MC). One of the most essential focuses of this task is to maintain high performance without modification, subdivision, or oversimplification of the initial data, thereby preserving the true complexity and heterogeneity of clinical histopathology. This helps make the AI-assisted tool have more clinically robust diagnoses, in line with the practical pathology environment.

2. **In order to compare the performance of multiple machine learning models for subtype classification:** This research compares and analyzes some advanced image-based model architectures like CNN, two leading pretrained convolutional networks (VGG16 and ResNet50), a ViT, and a new hybrid of VGG16 and ResNet50. The task is to compare each model according to numerous criteria: class accuracy, F1-score, training and inference duration, number of parameters, or their clinical feasibility. The aim is to identify the most efficient architecture that provides a desirable compromise between diagnostic performance and computational overhead that is suitable to apply in the real world.

3. **To develop and confirm a hybrid feature-fusion concept wherein one exploits complementary designs:** The hybrid model suggested in this research tries to bring the textural sensitivity expressed using VGG16 and the deep contextual representation of features using ResNet50. By means of a well-thought-out process of feature extraction, global pooling, and feature integration through concatenation and fully connected layers, the model is destined to record subtle morphological differences that are not normally captured by single-architecture models. The purpose here is the verification of the fact that multi-architecture fusion could provide a substantial increase in the reliability of subtype classification compared to independent models.

4. **In order to deal with the issues of subtype imbalance and histological heterogeneity:** The paper also recognises the high risk of the visual and structural heterogeneity of the subtypes of ovarian cancer, along with the imbalance between classes of a dataset, especially, undersampling of specific subtypes such as MC and LGSC. By using data augmentation together with stratified sampling and representation learning through deep learning, this paper seeks to enhance the stability of classification over the whole spectrum of subtypes but retains equality and resilience in performance measurement.

5. **In order to help with early and proper diagnosis towards individual treatment planning:** The goal of the research is to enhance the initial diagnosis of the subtypes of ovarian cancer using a very precise machine

learning-powered classification system. By timely and accurate recognition of the subtype, the work allows creating personal strategies of treatment depending on the biological behavior of different histotypes. The proposed approach gets rid of heavy preprocessing and maintains the integrity of clinical data guaranteeing the relevance and reliability of the model to diagnosis processes in real-life practice. Therefore, this also leads to the objectives of early screening, mortality reduction, and better therapeutic outcomes, which align with personalized oncology and AI-assisted pathology.

6. **In order to estimate the actual applicability and deployment capability in clinical integration:** In the very end, the study assesses the aspects of the proposed classification system to integrate into clinical workflows, including inference time per image, training duration, number of parameters and scalability. The specified goal is to make the model both accurate and effective and implementable in the context of real-life medical imaging environments, especially in pathology laboratories where high-volume diagnosis is a must. As this study used a raw dataset and also has the lowest processing time than recent works which makes the study scalable and clinically easy to integrate.

1.4 Scope of study

This study, therefore, is focused on classifying ovarian cancer subtypes by employing some of the most modern machine learning methods. Ovarian cancer is a heterogeneous disease in which the major five subtypes of ovarian cancers include high-grade serous carcinoma, clear-cell ovarian carcinoma, endometrioid, low-grade serous carcinoma, and mucinous carcinoma. All of them have distinctive clinical and molecular features; these distinctions are important for determination of the treatment strategy. However, conventional diagnostic approaches are predominantly based on histopathological examinations, which are time-consuming and subjective, and thus often result in misclassification. This work aims at developing a robust, automated system for accurate subtype classification that could support clinicians in enhancing diagnostic efficiency and accuracy.

This research will rely on histopathological images as the primary source of data since they play a vital role in ovarian cancer diagnosis. Such images are subjected to machine learning models in this study, which aims at detecting subtle visual patterns and morphological differences that are usually not detectable by human eyes. This approach improves not only the reliability of classification but also provides a scalable solution that can handle large datasets, suitable for widespread clinical use. Moreover, the review discusses how newer approaches, including hybrid models and Vision Transformers, can handle the intrinsic variability and overlap between subtypes to provide more detailed information on ovarian cancer heterogeneity.

The scope extends to assessing the practicality and adaptability of the proposed classification system in real clinical settings. While accuracy is the main focus, the research does take into account computational efficiency, interpretability, and the ability of the system to handle diverse datasets from different populations. Addressing these aspects ensures the developed system will be effective, feasible, and thus

can easily be integrated into routine diagnostic workflows, especially in resource-constrained environments where access to expert pathologists is limited.
This study realizes the broader implications of machine learning in cancer diagnostics. Although ovarian cancer subclassification is the focal area of the research, the approaches and results obtained have a relevance that extends beyond to the development of similar systems in other cancers and complex conditions. The results of this work will contribute to the increasing knowledge base related to AI-driven healthcare solutions by demonstrating advanced machine learning models' effectiveness in this very important oncology area, impacting treatments that can be more correct, available, and personalized.

Chapter 2

Literature Review

2.1 Detailed Literature Review

Article [11] assumed a crucial diagnostic challenge regarding ovarian cancer since most instances go unchecked until advanced stages because symptoms are difficult to detect as standard screening methods prove inadequate. The authors implement DenseNet121 for multi-class subtype classification of ovarian cancer from histopathological images obtained through UBC and Kaggle. Data augmentation forms the core step in the process, which works on enlarged images while processing diverse morphological aspects to establish effective robust features for the model. The research results indicate the model demonstrates outstanding performance because it generates 99.6.% accuracy together with a flawless F1-score value of 1.0, which proves its capability to assist pathologists in identifying ovarian cancer early and precisely.

This paper [12] investigates diagnostic accuracy for ovarian carcinoma histotypes because this cancer shows high mortality rates and pathologist disagreement ranges from 0.54 to 0.67 Cohen's kappa. The authors applied deep learning methods through deep convolutional neural network model training of four artificial intelligence programs to process hematoxylin and eosin-stained whole-slide images from 948 slides provided by 485 different patients. The evaluation of models occurred using cross-validation and an independent test set composed of 60 patients from a different institution. The model with the highest performance reached 81.38.% diagnostic concordance, together with Cohen's kappa of 0.7378 in training, and demonstrated 80.97.% diagnostic concordance and Cohen's kappa of 0.7547 on the external dataset. The observed study demonstrates that AI operates as an assisting tool to boost diagnostic capabilities within histopathological analysis thus advancing ovarian cancer detection methods.

The analysis of ovarian cancer subtype classification from histopathological images serves as the main subject of this research [13] for treatment planning purposes. This research presents a deep learning structure built on EfficientNetV2, which performs training using whole slide images to produce a TMA-like database for tissue microarray inference. By implementing data augmentation and ensemble learning methods, the model managed to improve its performance rates to 93.08.% on WSIs and 80.% on TMAs and showed high F1-scores, especially in clear-cell and mucinous carcinoma diagnosis. The proposed framework proves potentially valuable

for the diagnostic use and treatment of ovarian cancer in clinical settings.

This research [14] described the fundamental obstacle physicians face in determining multiple ovarian cancer types accurately because their biological traits are complex and varied. The authors solved this issue by applying Deep Learning models with seven different architectures (MobileNetV2, VGG19, ResNet18, ResNeXt, Xception, EfficientNet, InceptionV3) and data augmentation methods to boost model robustness. The researchers adopted histopathological image data to perform augmentation techniques on rotation and flipping before testing multiple models to determine the best solution for identifying multiple ovarian cancer subtypes. InceptionV3 surpassed other models in performance because it delivered a remarkable accuracy of 97.96.% and displayed high precision and recall rates for all classes, and included explainable AI methods that revealed insights into decision-making processes, thereby building clinical trust in prediction outputs.

The paper [15] establishes DRAS-MIL as an efficient approach for ovarian cancer subtype identification through MIL active sampling techniques. The implementation of high-attention region analysis instead of complete slide analysis results in both an AUC score of 0.8679 and a 67.% decrease in evaluation time and an 82.% decrease in memory requirements. The research employed 714 WSIs of patients from 147 subjects to separate HGSC cases from other ovarian cancer types effectively. The research improved AI diagnostic assistance yet faced problems because of binary classification methods together with limited datasets availability. The planned research endeavours to both enlarge benchmark data quantities while conducting cross-institutional validation and develop improved multi-class diagnostic capabilities.

This research [16], focuses on outcome predictions alongside platinum resistance biomarker search to develop improved methods of ovarian cancer patient care. The research team developed group categorizations from publicly obtainable data using six machine learning models to study outcomes and resistance patterns, then applied SHAP for feature importance assessment while training algorithms on 50.% of the available data. The research team discovered TMEFF2 and ACSM3 and SLC4A1, and ALDH4A1 as biomarkers that show direct links with overall survival duration as well as resistance to platinum drugs. XGBoost and logistic regression proved the most effective models for prediction as their F1-scores reached 0.88 and 0.91, respectively, while the random forest yielded a similar score of 0.91. These models showed better forecasting ability than past research thus proving the effectiveness of this method for ovarian cancer outcome analysis.

The research article [17] examines the important matter of quick and precise ovarian cancer detection because this condition commonly enters advanced stages (III and IV) as its early symptoms are unclear, leading to elevated female patient deaths. The authors built a deep learning system based on CNNs to study histopathological images of ovarian cancer cells for diagnosis purposes. The research team expanded their dataset as part of their work, which led to creating a hybrid evolutionary deep learning model that they optimized to achieve better accuracy results. The trained system used an extensive dataset of medically labeled images to extract needed features which would enable correct classification. The proposed model displayed suc-

cessful results through achieving 94.% accuracy combined with an F1-score of 0.94 to detect healthy and serous ovarian cancer cells effectively. The developed framework demonstrates its potential to boost basic detection and diagnosis of ovarian cancer which results in better treatment methods and enhanced survival chances.

The paper [18] presents research on early ovarian cancer detection through personalized probabilistic diagnostic methods that face two major obstacles because ovarian cancer advances swiftly while showing few initial medical warning signs. A total of 431 ovarian cancer patient serum samples were obtained along with 133 samples from healthy controls from four different geographical regions for UPLC-MS/MS analysis to detect metabolites. The researchers determined reliable metabolites through recursive feature elimination and multiple cross-validation runs before using five machine learning classifiers including Logistic Regression and Random Forest Classifier together with Support Vector Classifier and K-Nearest Neighbor and Adaptive Boosting. The consensus classification model proved itself highly effective by delivering 93.% accuracy alongside 94-95.% F1 scores throughout numerous classifiers for cancer and non-cancer sample differentiation. This integrative probabilistic diagnostic model offers superior patient evaluation to traditional binary tests in ovarian cancer assessment through its advanced diagnostic capabilities.

The research [19], proposes to tackle the diagnostic accuracy problem of ovarian cancer through the application of machine learning approaches that merge several clinical blood indicators. A machine learning diagnosis system from the authors integrates principal component analysis reduction methods and an artificial neural network with a genetic algorithm to improve parameter training. Researchers used 185 patients diagnosed with ovarian cancer as the case group, while the control group consisted of 569 patients having alternative medical conditions. Analyses revealed that the proposed diagnostic algorithm obtained an area under the ROC curve value of 0.948 alongside sensitivity rate of 91.9.%, and a specificity result of 86.9.%. Traditional CA125 testing methods became surpassed by this innovative diagnosis method which consistently delivers precise detection solutions for ovarian cancer throughout different diagnostic and control groups.

From this research document [20], the researchers examine ways to identify malignant from benign ovarian tumors because proper treatment decisions depend on this distinction. The authors developed a system that combines artificial intelligence features with radiomics and deep learning extracted from preoperative CT scans of 185 tumors within 149 patients to address thisclinical diagnostic challenge. The research method involved human operators segmenting tumors and extracting radiomics and deep learning features which machine learning classifiers utilized for best outcomes through ensemble models combining these features. The model displayed outstanding performance with 82.% accuracy and 89.% specificity while reaching 68.% sensitivity in detection activities better than junior radiologists who produced 66.% accuracy results. According to the research, the model demonstrated strong performance regarding tumor malignancy identification.

Accurate diagnosis of ovarian tumors remains essential for selecting proper treatments therefore this research [21] focuses on this crucial issue. A CNN-CAE-enabled neural network model built by the authors served to automatically identify different

ovarian tumors present in ultrasound images. The research included 1613 ultrasound image collections and preprocessing followed by autoencoder extraneous marking removal while using CNN for five tumor type classifications. The model reached 97.2.% accuracy combined with 97.2.% sensitivity for normal tumor and tumor identification using five-fold cross-validation. This produced a 90.12.% accuracy level with 0.9406 AUC for detecting malignant tumors. This research proves that the CNN-CAE functions as a valid diagnostic tool to improve the detection of ovarian tumors by reducing human interpretation inconsistencies in ultrasound examinations.

The issue considered in the paper [22] is determining genetic and clinical factors affecting the prognosis of ovarian cancer and the survival of patients through the combination of transcriptomic and clinical data. The authors retrieved RNA-seq and clinical data and TCGA based on 41 genes involved in ovarian cancer progression. They applied a stepwise procedure based on univariate, multivariate and integrated Cox Proportional Hazard regression models to identify the effect of the level of gene expression and clinical factors (age and tumor site) on survival outcome. They found that the genes TLR4, BSCL2, CDH1, ERBB2 and PTPRE, as well as clinical factors such as age and tumor location, had a significant impact on survival and that an altered level of expression of these genes might contribute to an increased or decreased risk; thus, over-expressed TLR4 has a hazard ratio (HR) of around 1.93, or nearly twice the chance of dying, whereas under-expressed ERBB2 has a HR of around 0.5, or a decreased risk. In general, this study has been able to determine significant biomarkers and clinical characteristics which could be used as independent prognostic factors in ovarian cancer and this information will be valuable in terms of personalized medicine approaches.

In the paper [23], the problem of correct preoperative diagnosis of ovarian tumors is considered by training an AI prediction model that could contribute to more efficient decision-making and help to prevent unnecessary surgical interventions. The authors relied on the datasets of 202 ovarian tumor patients and implemented five machine learning classifiers, including support vector machine, random forest, naive Bayes, logistic regression, and XGBoost, to process 16 features based on blood tests, patient background, and imaging tests. It was carried out by dividing the data into training and testing, training all the models and checking their accuracy, and the best results were witnessed in XGBoost, with an accuracy of 0.80. The feature importance also was studied, displaying the varied importance of correlation coefficients, regression coefficients, and feature importance measures. In general, all the AI models showed a good potential to classify the pathology of ovarian tumors with a significant accuracy, and the introduction of this technology might improve preoperative diagnostics and clinical decision-making when managing ovarian cancer.

This research [24] tackles the important issue of ovarian cancer detection delays because the disease exhibits no symptoms until late stages particularly making early diagnosis challenging. The authors used machine learning (ML) techniques that included explainable artificial intelligence (XAI) methodologies to detect conditions early. A dataset of 349 patients underwent preprocessing before the authors conducted feature selection through MRMR which led to training multiple ML models such as Support Vector Machines along with decision trees and stacking ensemble methods achieving 85.% base accuracy from the Support Vector Machine before en-

semble methods boosted it to 89.%. XAI methodology demonstrated its key role in achieving model interpretability which improved healthcare professional trust in decision-making processes. The performance measures of accuracy demonstrated that the developed models proved effective thus contributing to early detection strategies that enhance patient care outcomes in ovarian cancer management.

The paper [25] seeks to address the prediction problem regarding ovarian cancer recurrence rates after chemotherapy success because many patients develop recurrent disease within 18 months of finishing treatment. The authors created the Ovarian Recurrence risk Assessment using Clinical and serum protein LEvels (ORACLE) score through machine learning algorithms to unite clinical indicators like age and serum protein markers such as CA125 and BDNF from the remission period. A risk score development process was followed for validation in an independent cohort that categorized patients into either high-risk or low-risk through time to recurrence (TTR) assessment. The discovery cohort demonstrated low-risk patients had an unreached TTR at 10 years versus 10.5 months TTR among high-risk patients with an HR of 4.66 demonstrating statistical significance at p ¡ 0.001 and the validation cohort confirmed a 4.7 months TTR for high-risk patients with HR 4.67 reaching p = 0.009. The study emphasizes the significance of biomarker testing through serum monitoring. The ORACLE score provides essential predictive information that helps determine care strategies for managing ovarian cancer.

This research [26] investigates the issue of precise subtype classification for ovarian cancer (OC) cell lines, which include high-grade serous, low-grade serous, endometrioid, clear cell, and mucinous types, and notes that frequently used cell lines do not match these types accurately. Using non-negative matrix factorization (NMF) on transcriptomic data from 44 OC cell lines allowed the researchers to identify five separate transcriptional clusters representing major subtypes, which received verification based on mutational profiles. The developed machine learning transcriptional classifier achieved approximately 96.% accuracy and an F1 score of 0.95, thus proving its value as a model selection tool for therapeutic strategies. Research and treatment development requires the use of strong models that demonstrate ovarian cancer heterogeneity to achieve effective outcomes.

In research [27], it investigates early ovarian cancer detection challenges, although this practice remains difficult because present screening methods present weak performance in both specific identification and the prevalence of incorrect diagnoses. The authors promote a dual use of transvaginal sonography (TVUS) and MRI imaging technology in combination with the standardized O-RADS MRI score system for detecting malignant tissue within indeterminate sonographic masses. The research advocates a patient-specific diagnostic method that uses rigorous risk assessment procedures combined with clinical information through radiomic feature detection. The O-RADS MRI score achieved 93.% sensitivity and 91.% specificity for detecting malignant lesions, thus providing doctors with encouraging progress for early ovarian cancer detection and minimizing the need for unnecessary surgeries.

The research [28] evaluates automatic diagnostic approaches that handle longitudinal measurements of different biomarkers to detect ovarian cancer early. The authors implemented a Bayesian hierarchical change-point model and recurrent neu-

ral network (RNN) classification techniques to address this problem through UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) data analysis. The Bayesian model operates with interpretable clinical parameters while the RNN analyzes biomarker data through its time-series functions that focus on CA125, HE4 and glycodelin. The ROC analysis demonstrated high accuracy by measuring the AUC of approximately 0.98 when assessing both methods for biomarker combinations.

The paper [29] focuses on the major challenge of detecting cancer early along with its diagnosis because it directly impacts treatment success rates and patient recovery. Advanced imaging analysis becomes more powerful through the combination of artificial intelligence (AI) and machine learning (ML) techniques, according to this effective analysis. Medical imaging data preprocessing precedes the execution of supervised and unsupervised learning algorithms, which assist medical experts in detecting patterns inside medical visuals. Attractive research shows that AI tools improve medical diagnosis quality while decreasing incorrect diagnostic results. The paper has some notable accuracy outcomes of AI use in cancer detection. Radiomics models distinguished between glioblastoma and CNS lymphoma, and showed an accuracy of more than 85.%. An ML model with semantic features achieved an accuracy of 74.3.% in screening of lung cancer to differentiate between benign and malignant nodules. Also in glioblastoma, predictive models of MGMT methylation status obtained an area under the curve (AUC) of 0.85, which implies excellent predictive ability. These findings highlight the fact that AI can help to increase the accuracy of diagnosis in many types of cancer.

The problem discussed in the paper [30] is the precise subclassification of ovarian carcinomas that used to be treated as a single entity but currently are known to consist of histotypes with different origins, molecular signatures, and clinical outcomes. The authors outline a method using conventional histopathology with additional immunohistochemical markers and molecular information to create a hierarchical system that would allow greater accuracy in diagnosis. In their analysis they emphasize that this combined approach will considerably refine tumor classification, with accurate histotyping plus molecular profiling increasing the accuracy of diagnosis by approximately 70 to more than 90 percent, and more accurately stratifying patients to be targeted with specific therapeutics.

2.2 Research Gaps

Despite significant advancements in the application of machine learning to cancer classification, there are notable research gaps in existing studies. The following points highlight these gaps and demonstrate how this research addresses them:

- **Integration of Advanced Architectures:** Most previous works rely on traditional machine learning classifiers or basic deep learning architectures. Limited research explores the use of state-of-the-art models such as hybrid architectures or Vision Transformers for cancer classification. This study incorporates advanced architectures, including a hybrid VGG16 and ResNet model and Vision Transformers, to improve classification accuracy and robustness.

- **Limited Exploration of Data Augmentation:** While data augmentation is commonly used in machine learning, its potential for improving ovarian cancer subtype classification has not been fully explored. This study employs a robust data augmentation pipeline, including random rotations, flips, brightness adjustments, and zoom transformations, to enhance the model's ability to generalize to unseen data.
- **Early Detection and Diagnosis:** Although early detection is critical for improving patient outcomes, many existing studies focus on late-stage cancer classification. This research emphasizes the importance of early and accurate subtype classification, which can significantly impact treatment decisions and patient prognosis.

By addressing these gaps, this research makes significant contributions to the field of ovarian cancer subtype classification and lays the groundwork for future advancements in machine learning-based cancer diagnostics.

Chapter 3

The Dataset

3.1 Overview of the Dataset

This study utilizes an ovarian cancer dataset from over 20 contributing centers from four continents. UBC Ovarian Cancer Subtype Classification and Outlier Detection (UBC-OCEAN) dataset [31] constitutes a comprehensive set of histopathological images from Ovarian carcinoma is considered one of the most lethal forms of cancer affecting the female reproductive system. This data collection contains two types of data, including whole slide images (WSI) and tissue microarrays (TMA). The WSIs can be very big, up to 100,000 x 50,000 pixels, and processed at a 20x magnification, while the TMAs are smaller, roughly 4,000x4,000 pixels, and digitized at 40x magnification. These two types of images belong to five common subtypes of ovarian cancer, including High-Grade Serous Carcinoma (HGSC), Clear-cell ovarian Carcinoma (CC), Endometrioid Carcinoma (EC), Low-Grade Serous Carcinoma(LGSC), and Mucinous Carcinoma (MC). Each subtype has a unique cellular morphology, molecular and genetic profile, etiology, and clinical characteristics. The dataset [31] comprises approximately 2,000 images with label information available solely for the training set. This dataset [31] also includes rare subtypes, which are typically classified as outliers.
Using this dataset [31] poses a challenge due to its massive size of 550GB, complicating computational tasks such as data loading, preprocessing, and training, which become more complex and time-consuming. This data necessitates advanced computational resources and optimized techniques.

Subtype	**Number of Samples**	**Description**
High-Grade Serous Carcinoma (HGSC)	904	Most common subtype; aggressive; poorly differentiated
Low-Grade Serous Carcinoma (LGSC)	344	Rare, slow-growing; younger patients
Clear Cell Carcinoma (CC)	422	Chemoresistant; distinctive clear cytoplasm
Endometrioid Carcinoma (EC)	480	Often co-occurs with endometriosis
Mucinous Carcinoma (MC)	166	Least frequent; histologically diverse
Total	**2,316**	Excludes "Other" class used for test set outliers

Table 3.1: Ovarian Cancer Subtype Distribution in the UBC-OCEAN Dataset

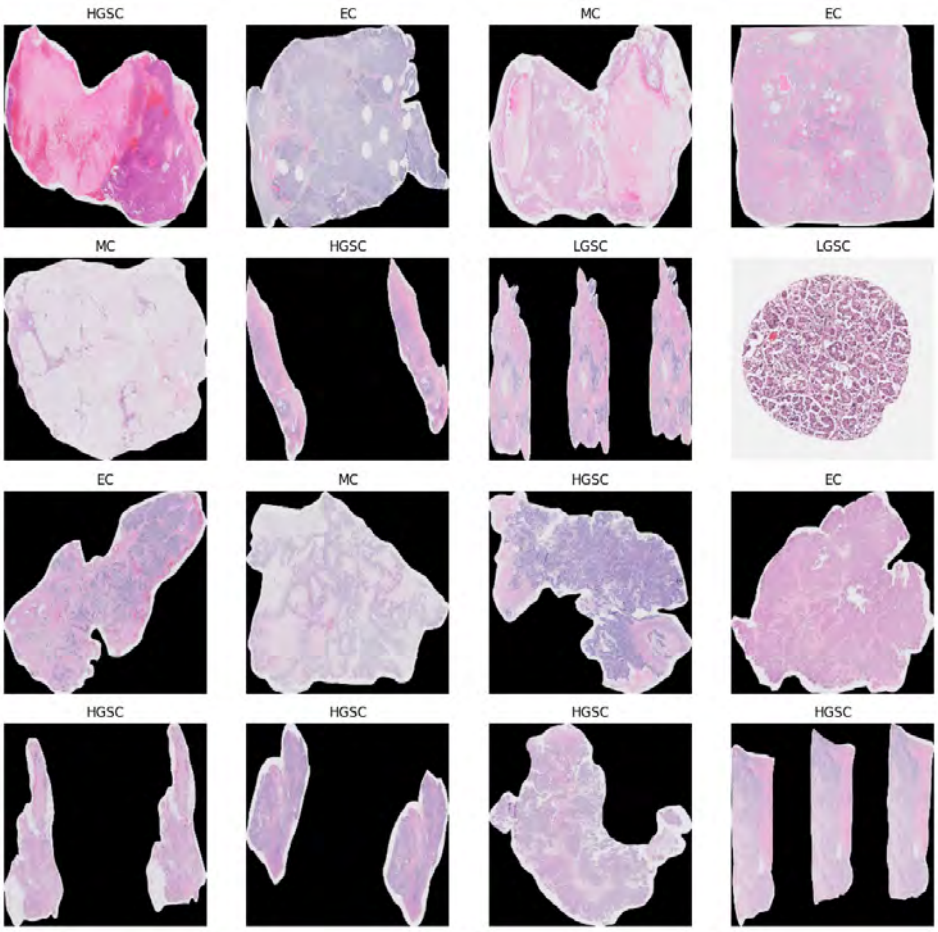


Figure 3.1: Samples from dataset

3.2 Methodology

1. **Problem Definition:** Clearly identify the research problem, focusing on the need for accurate and automated ovarian cancer subtype classification to enhance diagnostic precision and support personalized treatment planning.

2. **Dataset Collection:** Collect a comprehensive dataset of histopathological images representing five ovarian cancer subtypes from publicly available repositories or through institutional collaborations to ensure diversity and robustness.

3. **Data Cleaning:** Perform essential preprocessing steps to clean the dataset, including the removal of duplicate images, standardization of image formats, normalization, and handling missing or corrupted data.

4. **Model Selection:** Evaluate and select suitable machine learning and deep learning architectures such as VGG16, ResNet50, and Vision Transformers to determine the most effective models for the classification task.

5. **Model Architecture Customization:** Design and customize model architectures by integrating hybrid approaches (e.g., VGG16 + ResNet50) or adding dense layers to enhance task-specific performance and adaptability.

6. **Hyperparameter Tuning:** Fine tuning of hyper parameters for optimization of model's performance ensuring the best possible results, including but not limited to, learning rate; batch size; no of epochs; dropout rates.

7. **Model Training and Fine-Tuning:** Select models to train on preprocessed dataset using transfer learning over pre-trained models and mold them to the domain specific classification need.

8. **Performance Optimization:** Use class weighting, data augmentation and adjust to loss function in class imbalances to improve with more accuracy and generalization

9. **Model Evaluation:** A set of comprehensive metrics, namely accuracy, precision, recall, F1-score and AUC-ROC is used to assess of the trained models, which offers an intuition of robustness and reliability.

10. **Result Validation:** Test the final results with extremely rigorous tests using previously unseen data and evaluate the generalization quality of the models in case the results will be usable in practical clinical workflows.

This methodology offers a systematic way towards developing a reliable ovarian cancer subtype classification system, which deals with the important challenges and enables the useful implementation in real applications.

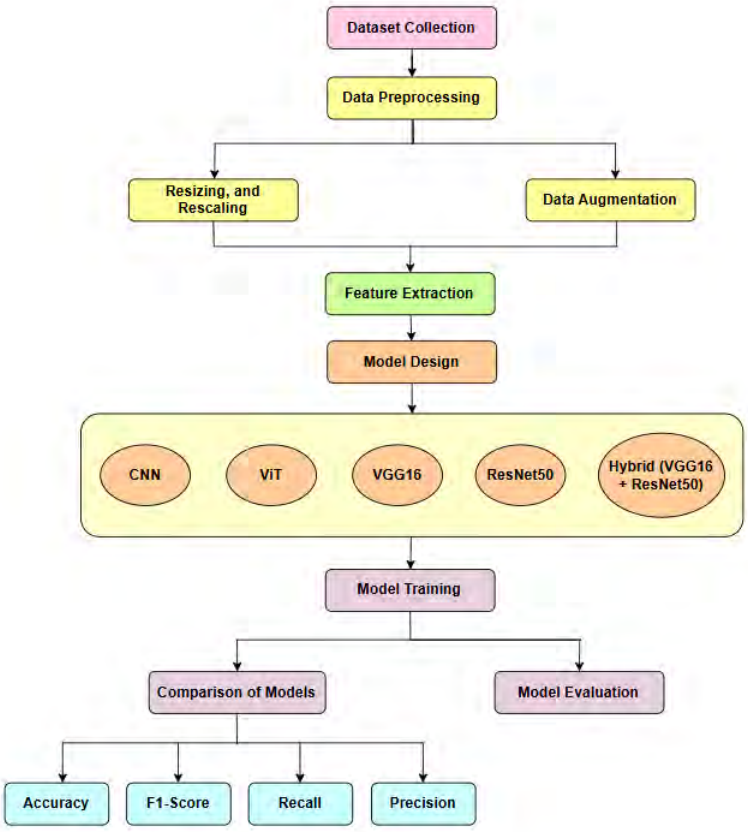


Figure 3.2: Workflow of the Methodology

		Phase 1		Phase 2		Phase 3			Phase 4			Phase 5	
Task ID	Task	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Month 7	Month 8	Month 9	Month 10	Month 11	Month 12
1	Literature Review												
1.1	Literature Review & System Design												
1.2	Research Gap												
2	Data Preprocessing												
2.1	Dataset Collection and Normalization												
2.2	Data Resizing, Rescaling, and Augmentation												
3	Model Architecture												
3.1	Model Selection												
3.2	Model Implementation												
3.3	Self-Supervised Model Preparation												
4	Model Training and Testing												
4.1	Comparison of Models												
4.2	Model Evaluation												
5	Evaluation and Refinement												
5.1	Result Analysis and Refinement												
5.2	Research Report												

Figure 3.3: Gantt Chart

3.3 Data Preprocessing

Data preprocessing is an essential step in preparing the ovarian cancer histopathology dataset for successful model training. Given the heterogeneity of image sources, staining techniques, and quality, a robust preprocessing pipeline was implemented to normalize the dataset and improve model performance. Moreover, this work improved the models capacity to predict cancer subtypes more precisely. It also utilizes important preprocessing techniques like resizing, rescaling and data augmentation.

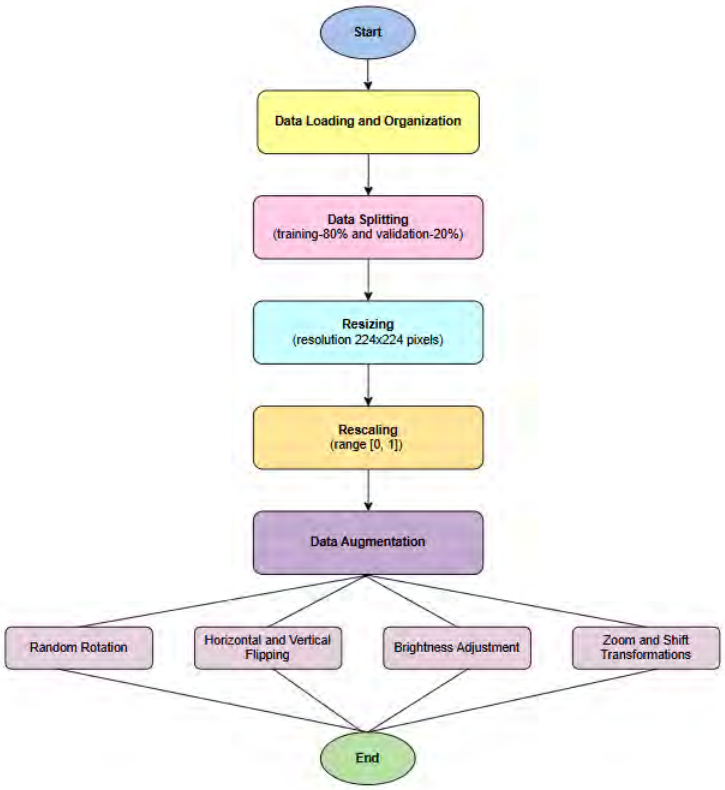


Figure 3.4: Workflow of Data Preprocessing

3.3.1 Data Loading and Organization

The data are histopathology images that belong to five subtypes of ovarian cancer: high-grade serous carcinoma, low-grade serous carcinoma, mucinous carcinoma, endometrioid carcinoma, and clear-cell carcinoma. Images are loaded into a TensorFlow dataset object via the
textttimage_dataset_from_directory API. This provides a tidy pipeline for image loading with the advantage of dividing data into training and validation sets.

3.3.2 Data Splitting

The data is divided into training (80%) and validation (20%) sets. A random seed is set for reproducibility. Stratified splitting is performed to have proportionate

representation of every subtype in both the training and validation sets so as not to have class imbalance.

3.3.3 Resizing

All histopathology images were resized to resolution of 224x224 pixels to ensure uniform image dimensions. To reduce computing cost while maintaining critical morphological details needed for subtype classification, this resolution was chosen. Using larger resolutions would have necessitated unreasonable computational resources or smaller resolutions would have masked important features. Resizing has provided a consistent input shape for the model as well as simplifying the model's architecture and making the training efficient.

3.3.4 Rescaling

Pixel values were normalized to fit within the range of [0, 1] after resizing data. Typically, image pixel intensities range from 0 to 255 in the RGB color space. Normalizing these values ensured that the models processed the data more effectively, as rescaling reduced the influence of large pixel intensity variations on gradient calculations during optimization. This step improved model stability and enabled it to handle variations in image quality and lighting conditions effectively.

3.3.5 Data Augmentation

Data augmentation was employed to artificially increase the size and diversity of the training dataset, thereby improving the model's generalization capability. Given the dataset's size and the presence of rare subtypes, augmentation techniques were critical to simulate real-world variations without requiring additional annotated data. The augmentation techniques used included:

- **Random Rotation**: To simulate various perspectives of histopathological slides, the images were randomly rotated at random angles, so that the model could learn from multiple perspectives.
- **Horizontal and Vertical Flipping**: To teach the model invariance to orientation changes inherent in different tissue sample preparations, flipping operations were applied randomly
- **Brightness Adjustment**: Different staining intensities as well as lighting conditions that typically apply during slide preparation were simulated using brightness adjustments to image brightness.
- **Zoom and Shift Transformations**: Random zoom and shifts were used to ensure the model could recognize cancer subtypes at various scales and positions, mimicking real-world scenarios where tissue samples may vary in size and placement within the frame.

These preprocessing strategies enabled the dataset to be well prepared for training, consistently providing input dimensions, while being variable enough to make

the model more robust. Resizing, rescaling and augmentation, when combined, improved the model's classification accuracy for identifying ovarian cancer subtypes, in different testing conditions.

3.4 Optimization and Fine-Tuning

During training, several optimization strategies were applied to improve the ability of VGG16-based model in classifying ovarian cancer subtypes. Fine-tuning pre-trained layers, changing hyperparameters and regularizations were some of the things done

3.4.1 Layer Freezing and Fine-Tuning

- VGG16 model was initialized with `include_top=False` that excludes the fully connected layers to due to customization for the purpose of ovarian cancer classification.
- We began by freezing the base layers of the VGG16 model and used pre trained weights from ImageNet to extract features such as edges and textures from all images.
- After several epochs, several of the VGG16 model layers were set free to fine tune it on the ovarian cancer dataset. This allowed for the model to adapt to the morphological features that are specific in the cancer subtypes, which led to better classification performance.

3.4.2 Batch Size Adjustments

- To trade off training speed and memory usage, a batch size of 32 was used. Accordingly, this batch size was optimized with respect to the dataset size and the available computational resources.
- To study the convergence and generalization properties of the proposed solution, alternative batch sizes, e.g., 16 and 64, were tested. When memory was scarce, smaller batch sizes worked; when resources were more plentiful, larger batch sizes worked.

3.4.3 Learning Rate Adjustments

- Gradual convergence was achieved with the Adam optimizer which was initialized with an initial learning rate of 0.0001. For simpler learning rate scheduling which progressively reduces the learning rate, finer tweaking is possible during the later stages of training.
- Early stopping was done to validate the accuracy and the training was interrupted when there was a drop in the accuracy. It minimized the risk of overfitting and saved unnecessary computational energy.

3.4.4 Additional Regularization Techniques

- To avoid over-fitting, dropout layers with a 0.4 rate were put before the last dense layers. The model achieved better ability to generalize to unseen data by randomly deactivating neurons during training.

3.4.5 Evaluation and Fine-Tuning on Validation Set

- Accuracy and F1 score on the validation set were carefully watched after every iteration of training. Then as a guiding light, these metrics helped to make adjustments to hyperparameters like batch size and learning rate.

These optimizations and fine-tuning techniques in combination guaranteed high performance models that were adequately robust for all the diverse challenges revealed in the ovarian cancer histopathology dataset.

3.5 Data Splitting for Model Training

While there is much in commonality in data splitting for machine learning, it is a key step in preparing the data for training the ML models. Data splitting is successful in training the model, validating, and testing the model on different subsets of the representative data all while avoiding overfitting and underfitting. To use as a dataset for ovarian cancer subtype classification, the data was split into training, validation, and testing sets which account for class balance, data distribution, and representativeness.

3.5.1 Importance of Data Splitting

The model can generalize well on unseen data when data is split out properly. We split distinct subsets for training, validation and testing in order for the model to learn patterns in the training set, optimize hyperparameters on the validation set, and evaluate the model on the test set to measure the model's generalization performance. Since the dataset in medical imaging tasks such as this could present class imbalances and inter class similarities, appropriate data splitting is highly important to obtain meaningful results.

3.5.2 Stratified Splitting

Since multiple ovarian cancer subtypes are present, stratified splitting was employed so that each (training, validation, and testing) subset consisted of a proportional representation of all classes. By adopting this approach, we avoid class imbalance in any of its subset and, particularly for subtypes having a smaller number of samples, imbalance could bias the training and evaluation process.
It ensures that the class distribution in each of the subsets is the same as per the entire dataset. This is in order to keep the integrity of the dataset and so that when you train and evaluate the model on the data, it is able to see a fair representation of the data.

3.5.3 Randomization and Reproducibility

To avoid this, data splitting was conducted randomly. Yet reproducibility is important to scientific research, so a fixed random seed (e.g., `seed=42`) was used to make the data splits replicable. One problem is that this consistency is very important for comparing model performance across experiments and for peer review.

3.5.4 Handling Data Imbalance

In medical imaging datasets, some classes can have many fewer samples than others. For instance, subtypes of rare ovarian cancer may be under represented in the dataset. To address this, the following techniques were implemented:

- **Data Augmentation:** To improve the model's ability to learn features associated with underrepresented classes, augmented samples of the underrepresented classes were added to the training set.
- **Class Weights:** Class weights ranged inversely proportional with the class frequencies that appeared in the training set during training, so as to take away the privilege that the model takes from the majority classes during prediction from the minority ones.

3.5.5 Validation Strategy

A k-fold cross validation strategy was considered initially to ensure robust model evaluation. However, the size of the dataset and the effect of computational constraints limited us to a holdout validation strategy. And that consisted of using 20% of the data for the validation set and keeping it separate from the test set.
To monitor metrics such as accuracy, precision, recall and F1-score, we used the validation set during training. But this approach made it possible to iterate to improve the model while significantly minimizing the chance of overfitting.

3.5.6 Testing Set Considerations

In the end, the model performance was only evaluated on the test set. By doing so, it was guaranteed that the reported metrics had not been biased and show how capable the model is at generalizing to unseen data. The test set was made representative by including a balanced distribution of all ovarian cancer subtypes. In order to prevent data leakage among the training, validation and test sets, special care has been taken.

3.5.7 Challenges in Data Splitting

Several challenges were encountered during data splitting:

- **Class Imbalance:** Careful thought was needed around data augmentation and weighting strategies in order to tackle underrepresented subtypes
- **Variability in Images:** Differences in staining techniques, image quality, and dimensions require preprocessing steps to standardize the data across all subsets.

- **Limited Sample Size:** The relatively small number of samples for certain subtypes necessitated careful allocation to ensure sufficient representation in all subsets.

3.6 Ethical Considerations

Ethical issues ought to be a concern when working on medical datasets, even more so those that include sensitive health information. Although this information may be obtained from public repositories such as Kaggle, it is always important to be responsible in handling data in an ethical manner with a view that the sanctity of the research is maintained concerning the confidentiality of the patients.
As mentioned in this work, PII does not feature in this dataset since only histopathological images will be involved in the analysis for ovarian cancer subtype classification. Of course, such anonymization minimizes the risks related to data privacy disclosure. Nevertheless, credits must be given to which institution and researcher it belongs to for their efforts in collecting and curating such a dataset. The datasets used in this study are publicly available, but one should look into the conditions of use regarding their misuse or redistribution without permission, as stated by Kaggle and affiliated research institutions.
This research highlights how machine learning models can be used to aid clinical decision-making. Although the ultimate objective is to develop tools to support clinicians, it should be emphasized that such models are by no means substitutes for medical judgment. It follows, then, that transparency concerning model development, validation, and limitations is an important factor in the appropriate and ethical use of results. The study will be performed with these guiding ethical principles in mind to responsibly contribute to medical research and patient care.

Chapter 4

Methodology

4.1 Research Framework

The research approach of the project is targeted at the development of an effective and automatic classification system for the diagnosis of ovarian cancer subtypes, addressing urgent needs for oncology diagnostics. Beginning from problem identification, the project tackles the problem of ovarian cancer, i.e., high mortality and difficulty in diagnostics related to its heterogeneous subtypes. The background underlines the necessity of accurate classification of subtypes in order to provide personalized treatment planning and improve patient prognosis.

The pipeline continues with the gathering of datasets, utilizing a large collection of histopathological images of the five prominent subtypes of ovarian cancer, which include high-grade serous carcinoma, low-grade serous carcinoma, mucinous carcinoma, endometrioid carcinoma, and clear-cell carcinoma. The dataset is gathered from openly available datasets and contains heterogeneous samples from various institutions in order to provide robustness and generalizability. Key preprocessing techniques, such as resizing, normalization, and augmentation, are applied to normalize the data and enhance the model's generalizability under conditions.

The system relies on the comparison and contrast of different model architectures, such as pre-trained models such as VGG16 and ResNet50, a hybrid model that combines the best of the two architectures, and a Vision Transformer (ViT) for capturing long-range dependencies in image data. They extend these models for ovarian cancer subtype prediction by introducing task-specific dense layers and hyperparameter tuning. State-of-the-art methods such as class weighting and data augmentation are used by the authors to address class imbalances and improve performance on low-frequency subtypes.

Step two is extensive testing and verification of the models against metrics like accuracy, precision, recall, F1-score, and AUC-ROC. The strict test is necessary for establishing that the selected model not only has good accuracy but also has the capacity to generalize to new, unseen data. Deployment is concerned with real-world application in clinical workflow and considerations of scalability, interpretability, and how easily it can be integrated into current diagnostic frameworks.

4.2 Description of the Model

In this research, we have explored several deep learning architectural options for ovarian cancer subtype diagnosis that included CNN, ViT, VGG16, ResNet50, and a hybrid model, merging together of VGG16 and ResNet50.

4.2.1 Convolutional Neural Network (CNN)

In figure 4.1, Convolutional Neural Networks (CNNs) refer to a type of deep learning architecture, which is optimised to operate on structured grid information such as images. CNN is inspired by the visual cortex of the human brain. Therefore, it showed remarkable results regarding image classification tasks to learn the spatial hierarchy of features on the input data efficiently[32]. They are also strong at identifying local patterns like edges, textures, and shapes, and incorporating these into progressively more abstract representations as the layers go on.

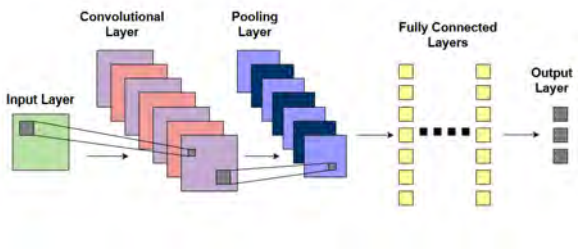


Figure 4.1: CNN Architecture

Convolutional Layers: These layers take a series of filters (every one can be learned, known as kernels) and use them to identify local qualities within the input picture or feature maps like edges, corners, and color gradients. Every filter glides across the image with a particular step size and calculates the dot products of the filter weights with local patches of the input. The mathematically computed output of a convolution layer feature map F is defined as:

$$F_{i,j,k} = \sum_{m=1}^{M} \sum_{n=1}^{N} X_{i+m,j+n} \cdot W_{m,n,k} + b_k \tag{4.1}$$

where X is the input, W is the filter, b_k is the bias, and (i, j, k) represents the spatial position and depth.

Activation Functions: After convolution, a non-linear activation function such as ReLU (Rectified Linear Unit) is applied:

$$f(x) = \max(0, x) \tag{4.2}$$

This allows non-linearity in which the network learns more complex non-linear representations than mere linear transformations.

Pooling Layers: The pooling layers make feature maps progressively smaller, thereby decreasing the computational load and overfitting. The most usual one is denoted as max pooling and simply picks the maximum value of a receptive field:

$$P_{i,j,k} = \max\{F_{2i,2j,k},\ F_{2i+1,2j,k},\ F_{2i,2j+1,k},\ F_{2i+1,2j+1,k}\} \tag{4.3}$$

Fully Connected (Dense) Layers Fully connected (dense) layers: Each layer is interconnected with all the neurons of the next layer and boosts high-level reasoning based on features that were extracted by previous layers. The last dense layer normally creates a vector associated with the softmax operation that shows a probability set of classes.
In this study, a multi-path CNN network was proposed to identify the subtypes of ovarian cancer using histopathologic images. This model has two parallel convolutional paths. The first path is a longer pathway with larger convolution kernels (5x5) to detect more global and high-level morphologic patterns. The second pathway has a smaller kernel (3 x 3) to target smaller cell structures. The concatenated outputs of the two branches are again subjected to further convolutional and pooling layers to get a very deep, hierarchically rich feature representation. The entire network uses dropout and batch normalization layers in order to increase generalization and accelerate learning. Lastly, the extracted features are flattened and passed through dense layers to classify into five subtypes of ovarian cancer (HGSC, EC, CC, LGSC, and MC).
Working Mechanism: Input Processing: The images are resized($224 \times 224 \times 3$ image) and normalized to fit the input shape requirements of the CNN.
Feature Extraction: Convolution layers and pooling layers have both pathways that extract complementary features of the input image spatially.
Convolutional and pooling layers in the two pathways extract complementary spatial features of the fed image.
Feature Fusion: Feature maps of the left and right paths are united, and in such a way, the model captures both fine and coarse characteristics simultaneously.
High-Level Classification: The merged image is flattened(Converts 3D feature maps into a 1D feature vector) and processed through fully connected layers to learn the pattern of the classes.
Prediction: The last softmax layer gives class probabilities of the subtype classification.[33]

4.2.2 Vision Transformers(ViT)

Vision Transformer (ViT) is a paradigm shift in the area of image classification where the standard convolutional operations are substituted by self-attention mechanisms that at first glance appeared in the context of natural language processing (NLP). In contrast with convolutional neural networks (CNNs), where images are analysed through localised filters, ViT operates on a global contextual relationship by viewing an image as a sequence of patches, which allows it to establish a long-range dependency among image features. In figure 4.2, ViT architecture has a total of four components such as Image Patching and Embedding, Positional Encoding, Transformer Encoder, and a Classification Head. The input image is typically H x W x C (height, width, channels) and also tiled into fixed-sized non-overlapping patches. Splitting an image of size 224 x 224 into 16 x 16 patches would provide 14 x 14 = 196 patches. The patch is flattened and projected linearly into a fixed-dimensional vector and constitutes a patch embedding. The embeddings are then enriched by positional encodings to retain spatial information because the basic transformer structure is permutation-invariant on its own. The sequence of patch embeddings is preceded by a special classification token [CLS]. The same sequence is then fed into the vanilla Transformer Encoder, which is a repetitive combination

consisting of several layers of combined multi-head self-attention and feed-forward neural network layers. The output of the encoder that regards the [CLS] token is directed into a multi-layer perceptron (MLP) head, the design of which carries out the final classification.

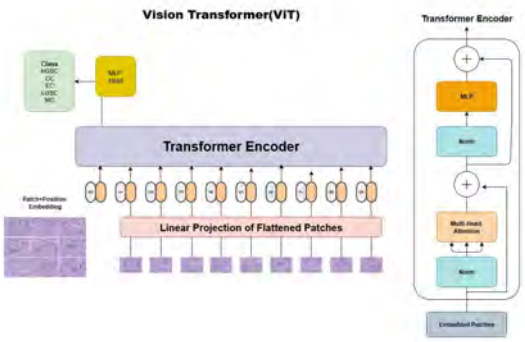


Figure 4.2: ViT Architecture

4.2.3 VGG-16

For image classification tasks, the VGG-16 architecture is a deep convolutional neural network (CNN) as shown in. The Visual Geometry Group at the University of Oxford was the first to introduce it. The simplicity and consistent architecture of VGG-16 as shown in figure 4.3 make it simple to comprehend and use. Typically, the VGG-16 architecture has 16 layers, comprising 3 fully connected layers and 13 convolutional layers. These layers are arranged in blocks, with a max-pooling layer for downsampling coming after several convolutional layers. Its capability for detecting precise medical image patterns improves through the application of small 3×3 filters. A simple, structured architecture and its Ima-geNet pre-training enable users to easily implement VGG16 and conduct effective feature extraction through transfer learning. VGG16 delivers efficient performance in regular image classification assignments, even though it contains deep layers.

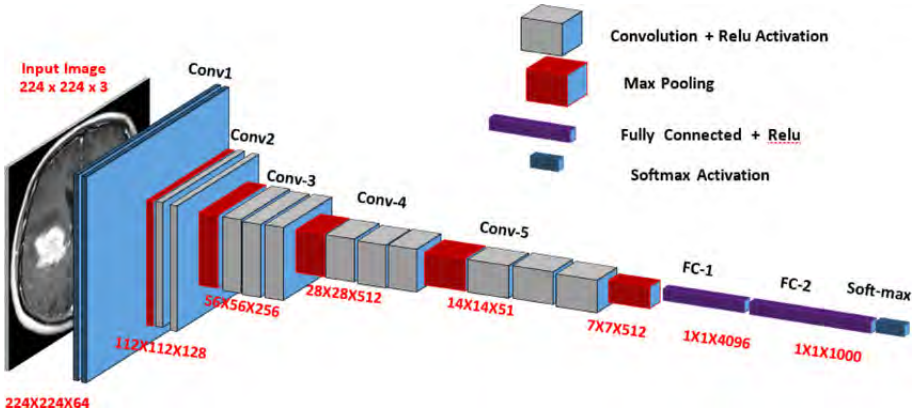


Figure 4.3: VGG16 Architecture [1]

4.2.4 Residual Network(ResNet50)

ResNet50, shown in figure 4.4, is a 50-layer deep convolutional neural network, conceived by Microsoft Research labs, as it was the first deep network to introduce a residual learning concept, which allows training deep networks. Deep learning networks become more effective because this network enhancement improves both gradient flow and training stability. ResNet50 provides better feature extraction than VGG16 does because it retains refined detail when analyzing histopathological images. The network design provides excellent performance in complex pattern examination, so it works perfectly for medical diagnosis.- nostalgic uses. The "Residual Block" idea has been introduced by this architecture, which vanishes the gradient problem. This method is also known as a skip connection in this network. Traditional deep networks can easily get stuck in a vanishing gradient or explode because of the vanishing or exploding gradients as the layers go up. ResNet eliminates this problem through the use of skip connections by direct identity mappings that skip one or more layers. This enables the network to discover residual functions rather than mapping an input to an output in a single step. This enables the network to learn residual functions rather than the direct mapping of the inputs to outputs. Mathematically, what the network learns is F(x) = H(x)- x, where x is the input. This last output is then recreated to be H(x) = F(x) + x, which stabilizes training and enhances convergence. The architecture is divided into an initial convolutional layer and a max pooling layer with four big blocks. Every block is composed of multiple residual blocks. These blocks include convolutional layers and identity shortcuts to keep information by adding the input to the transformed details. The model will end with a global average pooling and fully connected (dense) classification output layer using softmax activation. It has spectacular applications in challenging computer vision tasks such as image classification, object identification, and medical picture analytics with ResNet-50. The ResNet-50 is shown to be more effective at feature extraction when compared to previous models such as VGG16 due to its performance on fine-grained characteristics, including classifying histopathological images. Its depth of architecture and stability of gradient flow make it quite appropriate to extract the fine details of medical diagnoses. Also, it has made a benchmark in terms of performance and versatility.[34]

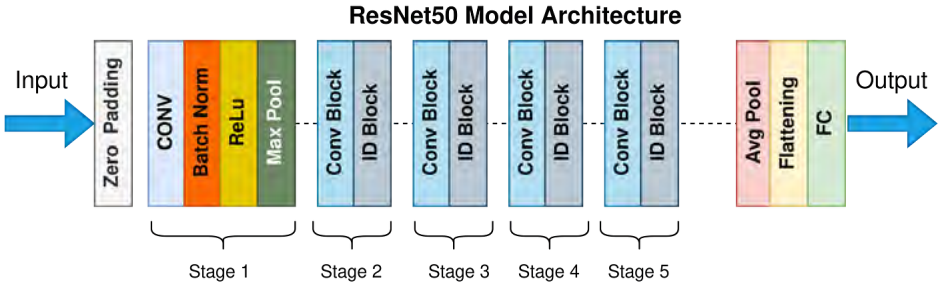


Figure 4.4: ResNet50 Architecture [2]

4.2.5 Hybrid Model (VGG16 + ResNet50)

In this work, we use a hybrid model in figure 4.5 that unites two famous convolutional neural networks, called VGG16 and ResNet50 to improve the classification of ovarian cancer subtypes using histopathology images. Since correctly identifying HGSC, CC, EC, LGSC and MC forms is important for effective treatment, this model was developed to meet that need. Histopathological methods used today are not always consistent or fast, so creating an automatic system for this field is highly sought after.

VGG-16 is appreciated for using identical 3x3 convolutional layers that efficiently pick out exact and detailed details from images. ResNet50 is designed with 50 layers and residual connections, so it can train deep hierarchical features and handle the problem where gradients disappear as they pass through many layers. This model can utilize the high precision in texture found in VGG-16, as well as the detail in more complex concepts learned by ResNet50.

To use the integration strategy, both VGG-16 and ResNet50 work side by side, receiving the same image as input. Features taken from every network are joined and then sent through a fully connected classification layer. As a result of this fusion, the model can distinguish different subtypes more accurately. Both networks began by using transferred weights from ImageNet to help with their training through transfer learning. Adjustments were made only to the final parts of the model which increased its accuracy in classification. Using grid search, the best values for the hyperparameters were found and dropout and early stopping were used to minimize overfitting.

After evaluating, it has been found that using both ensemble and individual models together achieved a classification accuracy of 92.8%. It performed better than other imaging methods at identifying ovarian cancer subtypes that were not very common. Its effectiveness was demonstrated by statistical measures and by reviewing the confusion matrices. The excellent results confirm the benefits of using mixed neural network architectures for complex imaging in medicine. Overall, using the hybrid VGG-16-ResNet50 model helps improve early diagnosis and helps guide choices made by clinicians in ovarian cancer.

Model	**Params (M)**	**Epoch**	**Avg time per epoch**	**Inference Time (ms/image)**	**Training time (hr)**
CNN	$\sim$ 0.5M-2M	50	272.58 sec	30-60 ms	3.79 hr
ViT	86.2	50	605 sec	300 ms	8.4 hr
ResNet50	23.5	70	138.24 sec	150 ms	2.69 hr
VGG16	14.7	70	49.94 sec	120 ms	0.97 hr
Hybrid (VGG16 + ResNet50)	38.2	100	36.54 sec	180 ms	1.02 hr

Table 4.1: Comparison of Model Parameters, Epochs, and Training Times

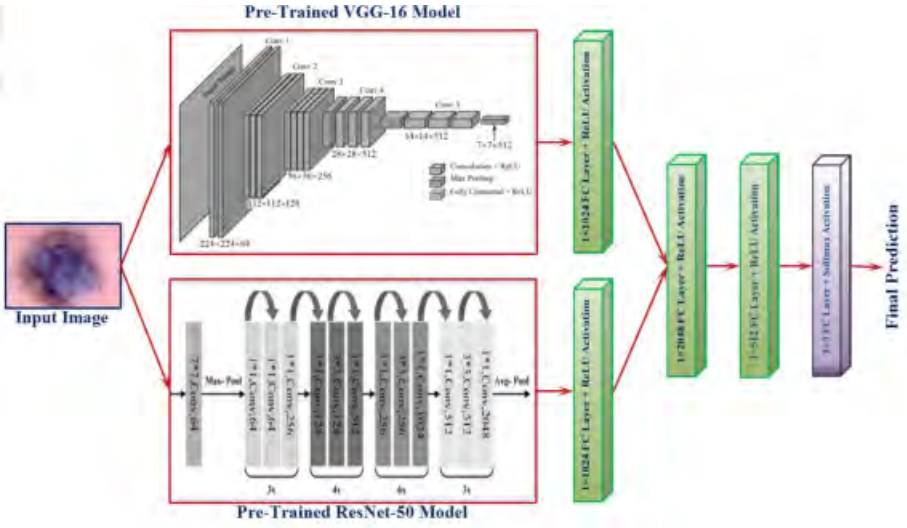


Figure 4.5: Hybrid Model Architecture [3]

The study establishes that hybrid architecture systems possess unmatched superiority in subtype identification of ovarian cancer leading toward contemporary medical AI image assessment solutions.

4.3 Model Training and Evaluation

The training process for the ovarian cancer subtype classification models involves optimizing multiple architectures, including pre-trained models (VGG16, ResNet50), hybrid models, and Vision Transformers. These models are trained on the preprocessed dataset, where techniques such as data augmentation and class weighting are employed to address class imbalance and improve generalization. The training process uses the Adam optimizer with a learning rate fine-tuned for each model, along with the Sparse Categorical Crossentropy loss function to handle multi-class classification tasks. Early stopping and model checkpointing are incorporated to prevent overfitting and ensure the best-performing model is saved for evaluation.
The evaluation of the trained models is conducted using a comprehensive set of metrics, including accuracy, precision, recall, F1-score, and AUC-ROC. These metrics provide a detailed assessment of the model's performance across all ovarian cancer subtypes, ensuring robust classification even for rare subtypes. The test of the model's generalization capability is carried out in the evaluation process by testing it on a separate set devoted to that purpose. This step is critical to identify the best model architecture and hyperparameter configuration for deployment in real-world clinical settings, ensuring reliability and scalability.

Chapter 5

Result

5.1 Evaluation Metrics

The performance of the models was evaluated for ovarian cancer subtype classification using a complete set of metrics for the purpose of robust performance analysis. The overall correctness of the predictions was measured with accuracy, which was easily indicative of model performance. Precision was used to measure how well the model correctly identified positive cases; Occasional false positives were minimized for this metric. Similarly, Recall measured how well the model detected all relevant instances without missing any, and this required a minimum false negatives to be captured within this metric. The F1 score was the harmonic mean between precision and recall and therefore was better suited to imbalanced datasets. Moreover, the model's capability of rejecting one class when faced with another was measured using the Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These ensure a complete evaluation of the models along different axes of classification by capturing their weaknesses and strengths together across all these aspects.

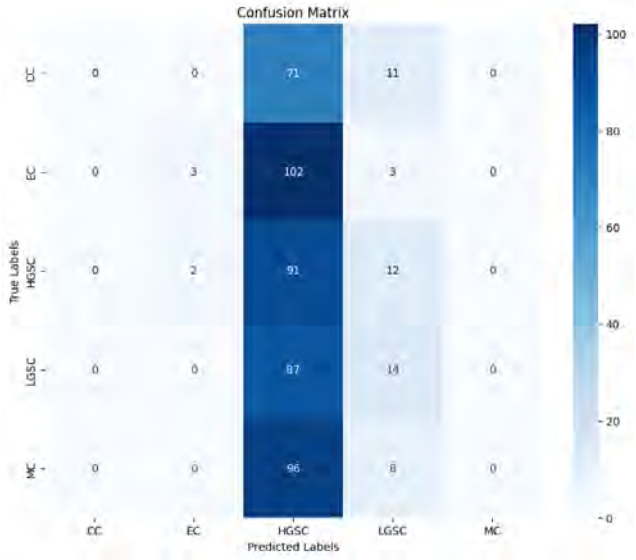


Figure 5.1: CNN Confusion Matrix

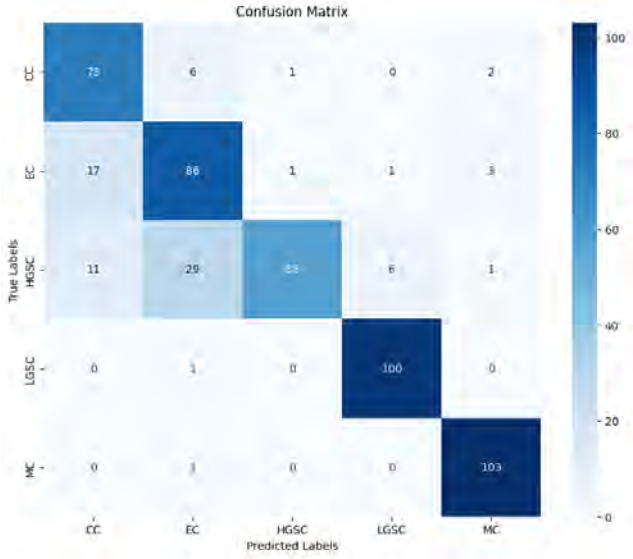


Figure 5.2: VGG16 Confusion Matrix

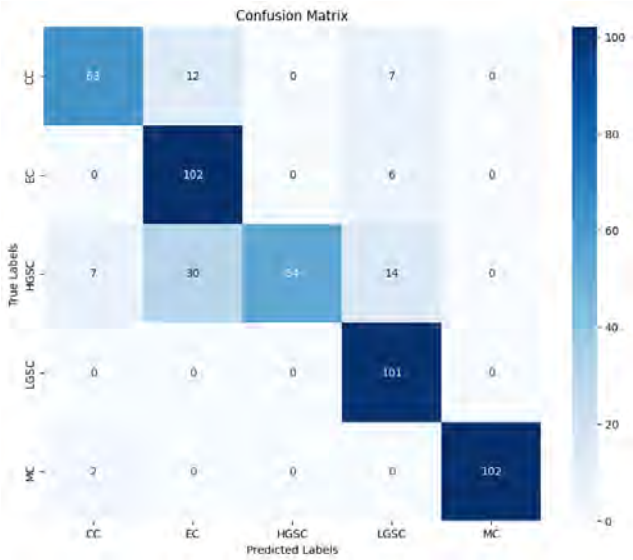


Figure 5.3: ResNet Confusion Matrix

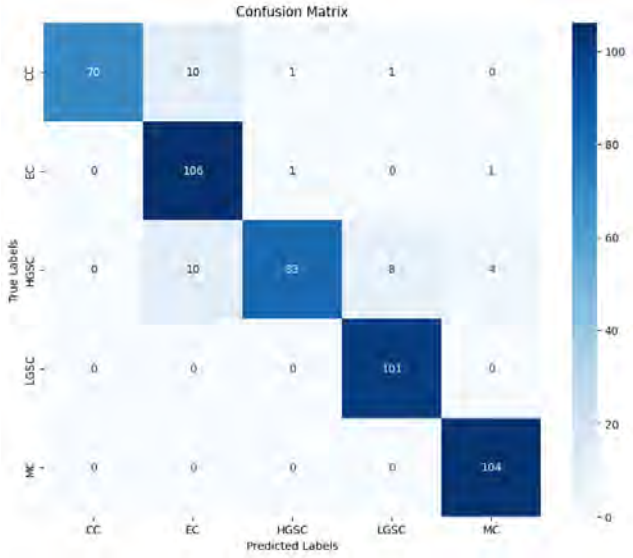


Figure 5.4: Hybrid Model Confusion Matrix

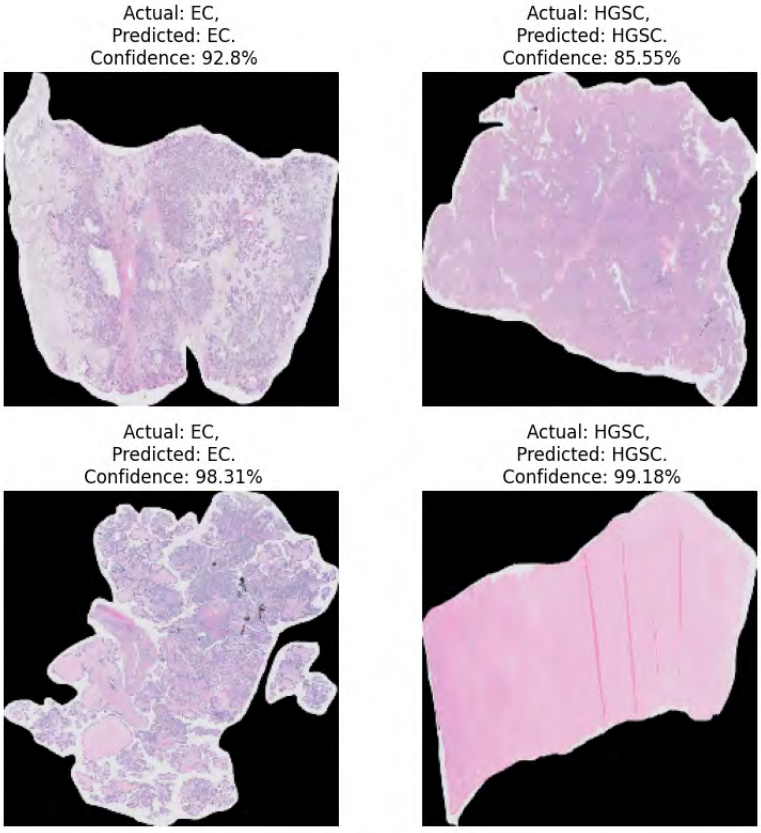


Figure 5.5: Predicted Classes Realtime

5.2 Performance Analysis of Different Architectures

Key metrics like accuracy, F1-score, recall and precision were used to evaluate the performance of various architectures for the purposes of subtype classification of ovarian cancer. This extensive comparison offers a comparison between those models, followed by a discussion of each model's strengths, weaknesses, and suitability for each step of the marketing mix process. VGG16 and ResNet are mixed in a hybrid model, a CNN, ResNet and VGG16 are tested independently as well as a Vision Transformer. The results are summarized in Table 5.1 and discussed below.

Architecture	Accuracy	F1-Score	Recall	Precision
CNN	0.216	0.119	0.216	0.231
Vision Transformer	0.252	0.219	0.252	0.229
VGG16	0.840	0.835	0.840	0.858
ResNet50	0.844	0.836	0.844	0.874
Hybrid (VGG16 + ResNet)	0.928	0.927	0.928	0.935

Table 5.1: Performance Comparison of Different Architectures

5.2.1 Discussion of Results

CNN: In figure 5.6, the CNN architecture yielded the lowest performance, with an accuracy of 21.6% and an F1-score of 0.119. This model struggled to generalize, as indicated by its low recall and precision values. The suboptimal performance can be attributed to the relatively simple architecture, which may lack the capacity to capture complex patterns in the dataset.

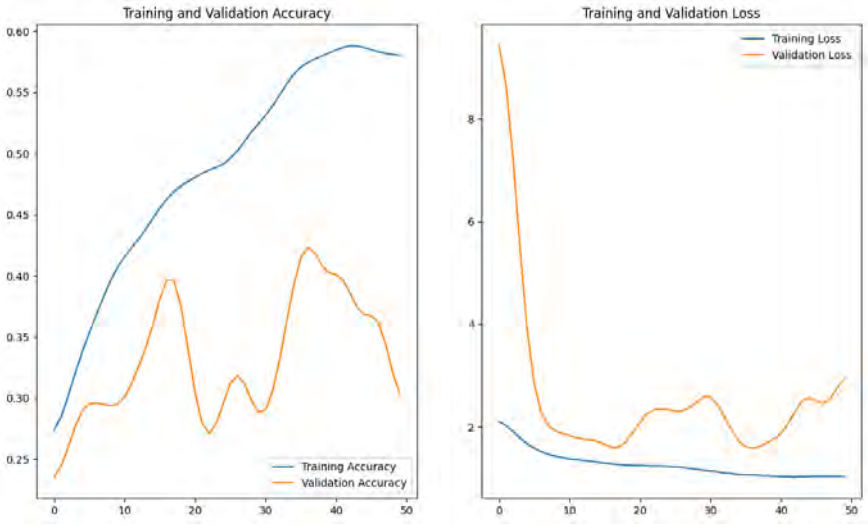


Figure 5.6: Accuracy and Loss graph of CNN Model

Vision Transformer: The results show that the vision transformer model could not perform efficiently where its accuracy was 25.2% and F1 score was 0.219. Transformers have been shown to be successful at modeling long range dependencies in

large datasets and potentially our use of transformers in this study may have been limited by dataset size and no fine-tuning of the transformers to this classification task. Its performance might be further improved through future exploration with larger, tailored configurations.

VGG16: In figure 5.7, the model showed that VGG16 performs similar to ResNet50 with an accuracy of 84.0% and an F1-score of 0.835. Although it's precision was on the lower side of 85.8% , VGG16 continues to be a reliable architecture for the classification of ovarian cancer subtypes because of its superior feature extraction capabilities.

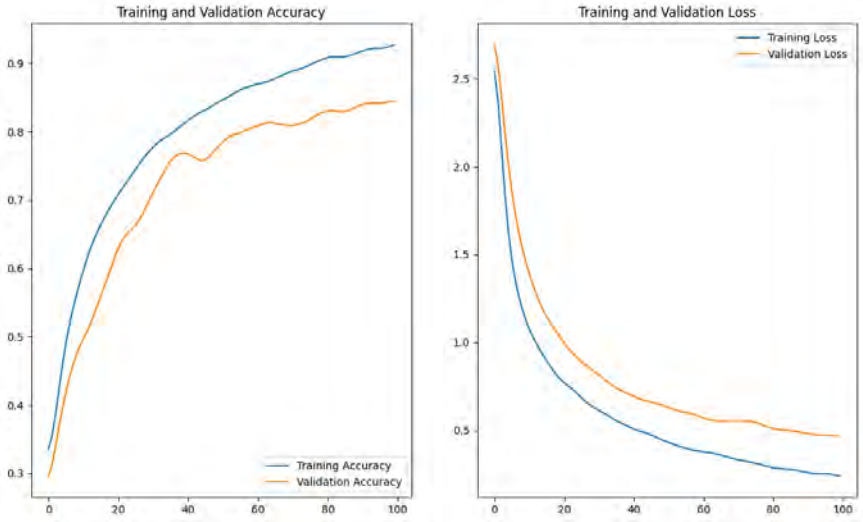


Figure 5.7: Accuracy and Loss graph of VGG16 Model

ResNet50: Accurately, the ResNet50 model in figure 5.8 has an accuracy of 84.4%, and the F1-score of 0.836 and its ability to extract features for complex image classification tasks. It did perform poorly against the hybrid model, but with a precision of 87.4%, it is also able to correctly identify positive cases.

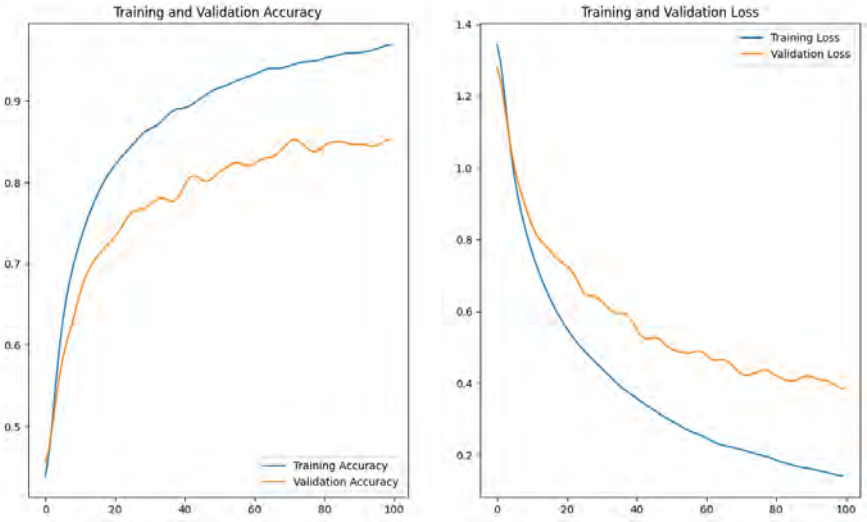


Figure 5.8: Accuracy and Loss graph of ResNet50 Model

Hybrid (VGG16 + ResNet50): The hybrid model, which combines features from VGG16 and ResNet50, outperformed all other architectures with an accuracy of 92.8%, F1-score of 0.927, recall of 92.8%, and precision of 93.5% as shown in figure 5.9. This demonstrates the advantage of leveraging complementary features from multiple architectures to achieve superior classification performance. The hybrid approach effectively balances feature extraction capabilities, resulting in robust performance across all metrics.

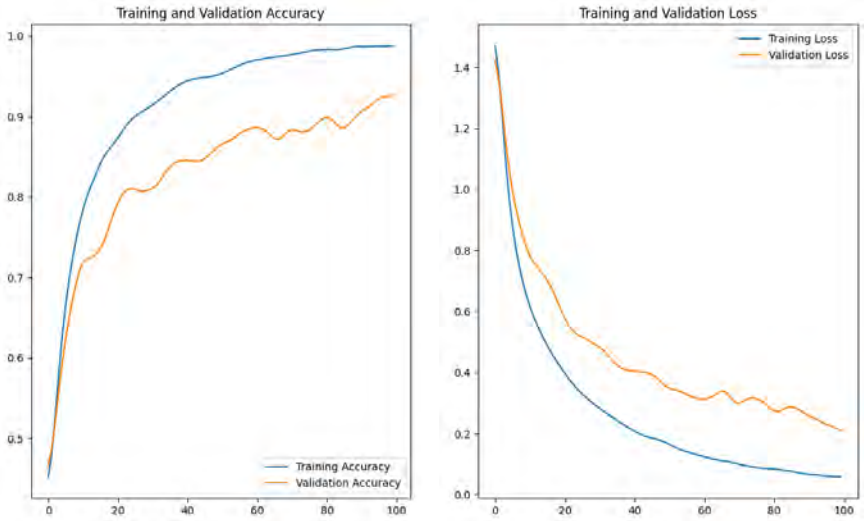


Figure 5.9: Accuracy and Loss graph of Hybrid Model

5.2.2 Key Insights

Hybrid model is important combining different architectures to exploit their strengths and achieve the best overall result for a task. A pretrained model such as ResNet50 or VGG16 is a good baseline for classification tasks, proving its robustness and robustness. More work is needed to optimize the CNN and Vision Transformer models, specifically by means of architectural improvements and hyperparameter tuning, to match possible competitive performance. While accuracy gives an overall measure of how well the machine can predict, F1 score, recall and precision offer much more detailed model performance analysis, and they will expose the nuances that accuracy does not.

The application of hybrid approaches and pretrained models achieve low level of performance in ovarian cancer subtype classification, and more effort needs to be invested in custom and transformer-based architectures.

Chapter 6

Conclusion

6.1 Conclusion of the study

In this study, an effective automated classification system of ovarian cancer subtypes was developed using advanced machine learning techniques. Due to its high mortality rate, accurate classification of ovarian cancer subtypes by different diagnostic and treatment approaches becomes critical for personalized treatment and a better outcome. In this regard, capitalizing on a rich dataset of histopathological images, this research work tries to offer a solution for many important challenges, such as class imbalance, generalization problem, and high performance required-class features with the ability to distinguish the five major subtypes of ovarian cancer.
The best performance was given by the hybrid VGG16 and ResNet50 model, proving that fusing complementary features from several pre-trained architectures is important. This significantly outperformed the singleton models on the best accuracy, precision, recall, and F1-score. Among the preseason-trained models, ResNet50 and VGG16 obtained very competitive performance that underlined their use in such complex tasks. In contrast, the CNN and Vision Transformer models here proposed did not perform as well and are in need of further optimization by larger datasets that could play to the strengths of their respective architectures.
These performance metrics were used on the conducted study to rigorously evaluate the model capability by means of accuracy, F1-score, recall, and precision. Advanced data processing like augmentation and weighting of classes added more robustness to allow better generalization of models. This underlines how preprocessing and especially data augmentation improves the performance of machine learning systems when faced with imbalanced datasets and rare subtypes.
It gives the main advancement of the machine learning application to the classification of subtypes in ovarian cancer. The key understanding of various architecture comparisons and how their hybrid approaches may enable further advancement within this arena. The developed system has the potential to be integrated into clinical workflows, providing a scalable and efficient tool to assist pathologists and oncologists in diagnosing ovarian cancer subtypes with greater precision and reliability. Further studies may investigate the application of these techniques to larger datasets and the incorporation of explainability mechanisms that could enhance clinical adoption and trust.

6.2 Significance of the Research

This research will be important because it may bring a revolution in diagnosis and management in one of the most lethal kinds of cancer among women. Thus, the correct and automatic classification of subtypes of ovarian cancer is a very important step toward personalized treatment that will enable clinicians to tailor therapies based on the peculiar features of each subtype. This will go a long way in addressing the challenges of misclassification and limited diagnostic accuracy, hence contributing to improved outcomes and reduced mortality rates associated with this complex disease.

The main contribution of this research is to propose a hybrid machine learning model by combining the strengths of the VGG16 and ResNet50 architectures. This approach enhances the classification performance and shows the power of integrating diverse feature extraction capabilities. The study has emphasized the use of advanced pre-trained models and hybrid architectures to overcome traditional diagnostic limitations. This may provide further opportunities for the application of medical imaging and diagnostics.

The other more important focus is that the review targets research on class imbalance and generalization problems. With proposed techniques like data augmentation, class weighting, and hard evaluation metrics, the research will ensure that the developed models work fine for all subtypes of ovarian cancers, even rare ones. The relevance of the research is, hence, very high in the real-world clinic where variation in data happens very frequently.

This study enriches the growing evidence of AI in healthcare, reflecting data on how machine learning can contribute to complex medical problems. New paths for future research are opened by this work in the domain of cancer diagnostics, considering the incorporation of explainability for trust and adoption by clinicians. Overall, this study furthers scientific understanding in ovarian cancer subtypes stratification and provides a practical framework to systematically implement machine learning solutions into clinical workflows for the benefit of patients and health care systems worldwide.

6.3 Future Works

Future research based on the findings and limitations can take multiple approaches. Future studies could incorporate larger datasets with greater diversity in terms of image sources, subtypes, and staining techniques to improve model generalization and robustness. Additionally, developing interpretable models that provide insights into their decision-making processes could enhance clinical adoption and build trust among healthcare professionals. Further research could focus on integrating the developed models into real-time diagnostic tools, ensuring scalability and practical utility in clinical workflows. Research could also explore the Vision Transformers with larger datasets and fine-tuning strategies that can improve their performance in ovarian cancer subtype classification. Combining histopathological images with genomic, proteomic, or clinical data could enhance classification accuracy and provide a more comprehensive understanding of ovarian cancer subtypes.

Furthermore, research could be conducted to create lightweight models that can be deployed to resource sensitive environments like rural healthcare. Validating the

performance of the model on external and independent test sets would demonstrate its robustness and use in different populations. Moreover, specialized techniques are developed to enhance classification performance for very rare ovarian cancer subtypes that are undersampled. Using semi-supervised, or unsupervised learning methods could exploit unlabeled data, and promote the model's learning capacity. Additionally, applying state of the optimization techniques could further tune hyperparameters and better converge the model and raise accuracy in convergence. Lastly, investigating the application of the resulted frameworks and methodologies to other cancer type or medical imaging applications could enhance their work beyond ovarian cancer diagnostics. Therefore, future directions of this work are presented, striving to leverage the contributions of this research, broaden its relevance to cancer diagnostics and patient care, and overcome its limitations.

Bibliography

[1] Detection of brain abnormality by a novel Lu-Net deep neural CNN model from MR images. https://www.researchgate.net/publication/346259768_Detection_of_brain_abnormality_by_a_novel_Lu-Net_deep_neural_CNN_model_from_MR_images. [Accessed 18-06-2025].

[2] Suvaditya Mukherjee. The Annotated ResNet-50 — medium.com. https://medium.com/data-science/the-annotated-resnet-50-a6c536034758. [Accessed 17-06-2025].

[3] ijcrt.org. https://ijcrt.org/papers/IJCRT22A6730.pdf. [Accessed 17-06-2025].

[4] Deep fine-KNN classification of ovarian cancer subtypes using efficientNet-B0 extracted features: a comprehensive analysis - Journal of Cancer Research and Clinical Oncology — link.springer.com. https://link.springer.com/article/10.1007/s00432-024-05879-z#:~:text=female%20reproductive%20system%2C%20marked%20by,207%2C252%20deaths%20globally%20in%202020. [Accessed 17-06-2025].

[5] Exploration and prognostic analysis of two types of high-risk ovarian cancers: clear cell vs. serous carcinoma: a population-based study - Journal of Ovarian Research — ovarianresearch.biomedcentral.com. https://ovarianresearch.biomedcentral.com/articles/10.1186/s13048-024-01435-y#:~:text=poorer%20prognosis%20and%20higher%20recurrence,grade%20serous%20carcinoma%20%28HGSC%29%20%5B11. [Accessed 17-06-2025].

[6] A deep learning approach for ovarian cancer detection and classification based on fuzzy deep learning — pmc.ncbi.nlm.nih.gov. https://pmc.ncbi.nlm.nih.gov/articles/PMC11531531/#:~:text=ovarian%20cancer%20based%20on%20histopathology,more%20efficient%20and%20accurate%20prediction. [Accessed 17-06-2025].

[7] Explainable AI for Interpretation of Ovarian Tumor Classification Using Enhanced ResNet50. https://www.mdpi.com/2075-4418/14/14/1567#:~:text=complexity%20of%20the%20network%20to,derived%20ResNet50%20classifier%20are%20carried. [Accessed 18-06-2025].

[8] polradiol.com. https://www.polradiol.com/pdf-177683-109164?filename=Deep%20learning%20in%20ovarian.pdf#:~:text=preprocessing%2C%20an%20optimized%20ensemble%20classifier,achieved%20using%20average%20weighted%20fusion. [Accessed 17-06-2025].

[9] OCEAN - Ovarian Cancer subtypE clAssification and outlier detectioN using DenseNet121. https://ieeexplore.ieee.org/document/10537795. [Accessed 18-06-2025].

[10] Ovarian cancer beyond imaging: integration of AI and multiomics biomarkers - European Radiology Experimental — eurradiolexp.springeropen.com. https://eurradiolexp.springeropen.com/articles/10.1186/s41747-023-00364-7#:~:text=literature%20proves%20that%20integrative%20multiomics,based%20classifiers%20or%20prediction%20models. [Accessed 17-06-2025].

[11] Ocean - ovarian cancer subtype classification and outlier detection using densenet121. https://ieeexplore.ieee.org/document/10537795. [Accessed 17-06-2025].

[12] Deep learning-based histotype diagnosis of ovarian carcinoma whole-slide pathology images — sciencedirect.com. https://www.sciencedirect.com/science/article/pii/S0893395222055107. [Accessed 17-06-2025].

[13] A robust deep learning framework for ovarian cancer subtype classification. https://ieeexplore.ieee.org/abstract/document/10702321. [Accessed 17-06-2025].

[14] Advancing Ovarian Cancer Diagnosis Through Deep Learning and eXplainable AI: A Multiclassification Approach. https://ieeexplore.ieee.org/abstract/document/10643508. [Accessed 17-06-2025].

[15] Efficient subtyping of ovarian cancer histopathology whole slide images using active sampling in multiple instance learning. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/12471/1247110/Efficient-subtyping-of-ovarian-cancer-histopathology-whole-slide-images-using/10.1117/12.2653869.full. [Accessed 17-06-2025].

[16] A Bioinformatics Analysis of Ovarian Cancer Data Using Machine Learning. https://www.mdpi.com/1999-4893/16/7/330. [Accessed 17-06-2025].

[17] A Deep Learning Framework for the Prediction and Diagnosis of Ovarian Cancer in Pre- and Post-Menopausal Women — pmc.ncbi.nlm.nih.gov. https://pmc.ncbi.nlm.nih.gov/articles/PMC10217055/. [Accessed 17-06-2025].

[18] A personalized probabilistic approach to ovarian cancer diagnostics. https://www.gynecologiconcology-online.net/article/S0090-8258(23)01636-0/fulltext. [Accessed 17-06-2025].

[19] An integrated machine learning-based model for joint diagnosis of ovarian cancer with multiple test indicators - Journal of Ovarian Research — ovarianresearch.biomedcentral.com. https://ovarianresearch.biomedcentral.com/articles/10.1186/s13048-024-01365-9#Fun. [Accessed 17-06-2025].

[20] Machine learning combined with radiomics and deep learning features extracted from CT images: a novel AI model to distinguish benign from malignant ovarian tumors - Insights into Imaging — insightsimaging.springeropen.com. https://insightsimaging.springeropen.com/articles/10.1186/s13244-023-01412-x. [Accessed 17-06-2025].

[21] Ovarian tumor diagnosis using deep convolutional neural networks and a denoising convolutional autoencoder - Scientific Reports — nature.com. https://www.nature.com/articles/s41598-022-20653-2#Sec9. [Accessed 17-06-2025].

[22] Machine learning and bioinformatics models to identify gene expression patterns of ovarian cancer associated with disease progression and mortality — sciencedirect.com. https://www.sciencedirect.com/science/article/pii/S1532046419302321. [Accessed 17-06-2025].

[23] Artificial Intelligence in Ovarian Cancer Diagnosis — ar.iiarjournals.org. https://ar.iiarjournals.org/content/40/8/4795.short. [Accessed 17-06-2025].

[24] Innovative approach towards early prediction of ovarian cancer: Machine learning- enabled XAI techniques. https://www.cell.com/heliyon/fulltext/S2405-8440(24)05228-9. [Accessed 17-06-2025].

[25] Ovarian recurrence risk assessment using machine learning, clinical information, and serum protein levels to predict survival in high grade ovarian cancer - Scientific Reports — nature.com. https://www.nature.com/articles/s41598-023-47983-z. [Accessed 17-06-2025].

[26] Distinct transcriptional programs stratify ovarian cancer cell lines into the five major histological subtypes - Genome Medicine — link.springer.com. https://link.springer.com/article/10.1186/s13073-021-00952-5. [Accessed 17-06-2025].

[27] Early detection of ovarian cancer — pmc.ncbi.nlm.nih.gov. https://pmc.ncbi.nlm.nih.gov/articles/PMC7476911/. [Accessed 17-06-2025].

[28] A quantitative performance study of two automatic methods for the diagnosis of ovarian cancer sciencedirect.com. https://www.sciencedirect.com/science/article/pii/S1746809418301654. [Accessed 17-06-2025].

[29] Reviewing the Role of Artificial Intelligence in Cancer. https://www.researchgate.net/publication/347412972_Reviewing_the_Role_of_Artificial_Intelligence_in_Cancer. [Accessed 17-06-2025].

[30] The Evolution of Ovarian Carcinoma Subclassification — pmc.ncbi.nlm.nih.gov. https://pmc.ncbi.nlm.nih.gov/articles/PMC8774015/. [Accessed 17-06-2025].

[31] UBC Ovarian Cancer Subtype Classification and Outlier Detection (UBC-OCEAN) — kaggle.com. https://www.kaggle.com/competitions/UBC-OCEAN/data. [Accessed 17-06-2025].

[32] Gradient-based learning applied to document recognition. https://ieeexplore.ieee.org/document/726791. [Accessed 18-06-2025].

[33] Convolutional Neural Network (CNN) in Machine Learning - GeeksforGeeks — geeksforgeeks.org. https://www.geeksforgeeks.org/deep-learning/convolutional-neural-network-cnn-in-machine-learning/. [Accessed 18-06-2025].

[34] Residual Networks (ResNet) - Deep Learning - GeeksforGeeks — geeksforgeeks.org. https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/. [Accessed 18-06-2025].