SPECTRUM OF ENGINEERING SCIENCES
ISSN (E) 3007-3138 (P) 3007-312X VOLUME 4, ISSUE 3, 2026

HARNESSING THE POWER OF ARTIFICIAL INTELLIGENCE FOR ACCURATE OVARIAN CANCER CLASSIFICATION

Saira Amjad[1], Saeed Ahmad[2], Muhammad Shadab Alam Hashmi[3*]

[1,3]*Institute of Computer Science, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan 64200, Pakistan*
[2]*Department on Computer Science, Govt. Khawaja Fareed Graduate College Rahim Yar Khan 64200, Higher Education Department (HED) Punjab, Pakistan*

[1]*sairaamjad747@gmail.com*, [2]*saeedcdkfcryk@gmail.com*, [3*]*shadab.alam@kfueit.edu.pk*

DOI:https://doi.org/10.5281/zenodo.22275650

Keywords

Article History
Received: 01 January 2026
Accepted: 17 March 2026
Published: 31 March 2026

Corresponding Author: *
Muhammad Shadab Alam Hashmi

Abstract
Ovarian cancer's high death rates, which are mostly brought on by late-stage detection, make it a major global health concern. Early detection is needed to improve survival and for effective treatment. This study explores how artificial intelligence (AI) might improve the identification of ovarian cancer, which is in line with Sustainable Development Goal 3 of the UN, which aims to promote health and well-being. We propose a combination of machine learning and transfer learning techniques, based on an open-source medical imaging dataset specifically for ovarian cancer. To extract intricate and hierarchical information from medical photos the study uses the pre-trained VVG16 deep learning model, The XGBoost algorithm, a potent ensemble learning method renowned for its excellent performance and efficiency was then used to classify these characteristics. The suggested approach showed robustness and reliability with an outstanding accuracy of 98.38% on the dataset in detecting ovarian cancer. The study points that there is great potential in the promise of AI-driven solutions and transfer learning in medical diagnostics, especially for diseases like ovarian cancer where early and accurate detection can greatly enhance results. The insights underscore the potential of AI to revolutionize healthcare and improve patient outcomes by providing timely treatments, potentially leading to the development of scalable, effective, and accurate diagnostic tools.

INTRODUCTION

Ovarian cancer is one of the critical health challenges around the world with a high mortality rate due to the complexity of diagnosis. Advanced stage epithelial ovarian cancer (EOC) initially seems to be susceptible to chemotherapy following cytoreductive surgery since platinum-based therapy response rates are higher than 80% (Bookman et al. 2009). However, due to recurrence and the development of treatment resistance, long term survival is still poor. In spite of the continued use of platinum-based therapies (like cisplatin or carboplatin) and Texans as the preferred first treatment, other cytotoxic drugs have been tried in clinical trials, including topotecan, gemcitabine, and methoxy polyethylene glycosylated liposomal doxorubicin (PLD). Each drug has different action in the setting of recurrent EOC, even in populations that are platinum resistant. Topotecan and PLD have been approved as single agents for recurrent EOC by the US Food and Drug Administration. High-

grade serous, low-grade serous, clear cell, endometriosis, and mucinous are the five histological subtypes of ovarian malignancies, which are mainly carcinomas that arise from epithelial cells (Breen et al. 2023). Non-epithelial ovarian malignancies, such as mesenchymal, germ cell and sex cord-stromal tumor, on the other hand, occur much less commonly. The prognosis, treatment and appearance of each subtype of ovarian cancer is different. The most common subtype accounting for around 70% of all ovarian cancer cases is high-grade serous carcinoma. Different tumor types with unique genetic alterations and behaviors are high-grade serous carcinomas (LGSCs). The most prevalent kind of ovarian cancer, HGSCs are frequently identified at an advanced stage and are linked to chromosomal instability, BRCA abnormalities, and TP53 alterations. LGSCs, on the other hand, are less common typically have a serious borderline component and have mutations in KRAS and BRAF. Although most cases involve ovarian cancer involvement, HGSCs can originate in the tubal or peritoneal regions (Prat and Gynecologic Oncology 2014).

The seventh most prevalent disease in women is ovarian cancer, especially epithelial ovarian carcinoma (EOC), although early detection might be difficult. Lower survival rates result from the majority of patients being diagnosed at an advanced stage. The EOCSA framework, which applies deep learning on whole-slide images (WSIs) for EOC survival analysis, is presented in this paper. Deep ConvAttentionSurv (DCAS), a deep survival model, is used by the framework to extract more useful features and increase prediction accuracy. The work proposes two novelties: improving feature extraction and weight calculation techniques and utilizing WSIs to apply deep learning to EOC survival analysis (Liu et al. 2022). This study investigates the association between tumor genetic markers and the predictive. This study analyzes how tissue specimens under the microscope help predict high-grade serous ovarian cancer outcomes (HGSOC). This approach recommends combining multiple dimensional features to improve prognostic predictions. A multi-dimensional integration framework unites transcriptomic data sets with histopathological image views. proteomics, and genomes. A prognostic model integrates predictions of mutations alongside MSI status and molecular subtypes. The development of a predictive model required machine learning algorithms to which data scientists applied before model verification. The Testing yielded improved prognosis prediction accuracy for patients sustaining HGSOC. A combination of multi-omics information with histopathological characteristics leads to performance enhancement (Zeng et al. 2021). The researchers propose a transfer learning approach as a method to boost ovarian cancer classification capabilities.

To enhance the classification of ovarian cancer, this study suggests a transfer learning strategy. The study assesses the model's performance on an open-source ovarian cancer dataset using the XGBoost classifier for classification and the pre-trained VGG16 model for feature extraction. The effectiveness of the suggested model is evaluated using performance metrics like accuracy, precision, recall, and F1-score. Comparisons with other pre-trained models, such as InceptionV3 and ResNet50, are also made. Through this study, early detection methods in ovarian cancer are advanced including the promotion of the diagnostic accuracy and classification categories that is associated with Sustainable Development Goal 3. This study is divided into four parts: conclusion, result and discussion, materials and procedures, related work.

Ovarian Cancer Causes

Ovarian cancer is a disease that develops when cells in the ovary undergo changes in their cells and genes that allow the cells to multiply rapidly and form a tumor. These changes could be induced by the body that may not be able to repair the damage to the DNA which can occur as a result of ageing and inflammation or exposure to toxins. When there is a ratio imbalance between the estrogen and progesterone hormones, this can promote abnormal growth of cancer cells in the ovary. Early malignant cells may go undetected and unrestrained in their proliferation if the immune system is unable to identify and eliminate them. Reactive oxygen species (ROS) which are created during cellular metabolism or persistent

inflammation can harm DNA and aid in the development of cancer. It is important to understand these causes and risk factors to develop ovarian cancer prevention plans and personalized treatment options for early detection. A mix of lifestyle environmental and hereditary variables can lead to ovarian cancer. Although the precise cause of the disease is frequently unknown, a number of important elements have a role in the development. The risk of ovarian cancer is greatly increased by mutations in these genes. These genes are in change of repairing damaged DNA, and when they malfunction genomic instability results which makes cancer possible. The risk of ovarian cancer is also increased by a genetic disorder brought on by mutations in the mismatch is the mismatch repair genes.

Women are more vulnerable if they have close family members (mother, sister or daughter) who have had colorectal, breast, or ovarian cancer. Most occurrences of ovarian cancer are detected after the age of 50, and they are more common in postmenopausal woman.

Long term exposure to estrogen puts women who have never been pregnant or who began menstruation early and went through a late menopause at higher risk. After menopause, using estrogen-only HRT for an extended period of time may raise risk. Being overweight is linked to an increased risk of ovarian cancer, especially high-grade serous ovarian cancer. Although the research is inconclusive, a diet that is low in fruits and vegetables and high in saturated fats may raise the risk. A particular subtype of ovarian cancer called mucinous ovarian cancer has been associated with smoking. Certain ovarian cancer subtypes are likely to occur in women who have endometrioses, a disorder in which tissue from the uterine lining develops outside the uterus. The risk of ovarian cancer may be somewhere increased by this hormonal condition. Persistent pelvic inflammation may increase the risk of developing cancer.

Over time, ovulation-induced chronic inflammation of the ovarian epithelium can damage DNA and raise the risk of developing cancer. Aggressive types of ovarian cancer, like high grade serous carcinoma (HGOEC), can result from differences within tumor cell populations. DNA methylation is one example of a shift in gene expression that can inactivate tumor-suppressor genes and encourage the growth of cancer without affecting the DNA sequence. Although this is still up for debate, some research points to a potential connection between talcum powder use in the vaginal area and ovarian cancer brought by asbestos exposure.

Literature Review

In this study, an innovative optical biopsy system based on deep learning technology is introduced that can be used in combination with SHG imaging to achieve precise and rapid surgical diagnosis of ovarian cancer. When the researchers tested the per-trained CNN models on SHG images, they achieved a 99.7% accuracy to distinguish between benign and malignant and normal ovarian cancer tissues. Surgical procedure in-the-moment diagnostic capabilities demonstrated in this method would lead to both improved medical results and reduced medical costs.

The work by Blessed Ziyambe et al (2023) introduces a recent deep learning technique to predict and diagnose ovarian cancer through histopathological imageries. Similarly, the utility of deep transfer learning in clinical oncology is demonstrated by the osteosarcoma bone tumor detection framework. This approach employs a deep ensemble transfer learning architecture to extract and classify critical tumor boundaries in radiographic medical imaging (nourin et al. 2025). The CNN model demonstrated high performance in both cancer detections with 94% accuracy and normal cell identification. The research addresses two key limitations of expert clinical examination by minimizing both expert misdiagnosis rates and observation variability. The framework provides an advanced method that shows potential to enhance diagnostic outcomes and treatment results for ovarian cancer patients (Ziyambe et al. 2023).

MUNETOSHI AKAZAWA (2020) utilized AI technology to estimate ovarian tumor pathological evaluation through assessment of patient and preoperative diagnostic results. The analysis of 202 patient data points using five machine learning classifiers demonstrated that XGBoost reached 80% accuracy. Feature importance revealed

albumin and ascites, and CRP showcased predictive potential. This study examined medicine's growing usage of AI technology together with the benefits this breakthrough offers to diagnostic precision. The study recognized both positive and negative effects resulting
from the available data set size. The collected research indicates artificial intelligence methods hold potential value for diagnosing ovarian tumors pathologically (Akazawa and Hashimoto 2020).

A predictive model for ovarian cancer was developed through machine learning applications focusing on Random Forest (RF). With RF achieving the highest level of accuracy and ROC curves' AUC rate for differentiating EOC from benign tumors (resulting in 92.4% accuracy, 0.968 AUC) and stage prediction with 69.0% accuracy, 0.760 AUC. The system accomplished both histology classification and surgical resection status
assessment. The research methodology distinguished between survival outcome groups which exist within early-stage EOC patients. Through pre-treatment stratification, this methodology permits individualized care through customized therapeutic strategies.

Guangxing Wang (2023) introduced an innovative surgical diagnosis method using optical biopsy techniques to detect ovarian cancer during operations swiftly and with precision. Researchers used deep learning techniques to analyze fresh unprocessed ovarian tissue samples during resection through second harmonic generation (SHG) imaging. Research employed a convolutional neural network framework based on ResNet50 with SHG image fine-tuning to differentiate between normal, benign, and malignant ovarian tissue types. Researchers collected 13,563 second harmonic generation images from 74 patients in the SHG data collection. The fine-tuned CNN system proved successful at identifying ovarian tissue types with 99.7% average accuracy thus validating the combination of deep learning with optical biopsies for fast and precise surgical ovarian cancer diagnosis (Wang et al.
2023).

This research by Aditya MS (2021) applies machine learning classifiers along with deep learning approaches to tackle the vital diagnosis distinction between Benign Ovarian Tumor and Ovarian Cancer. A dataset provided by Kaggle included information about 350 patients with their 50 respective healthcare parameters such as age and menopause status in addition to tumor type diagnosis. Model performance improved through feature selection together with imputation methods. The combination of Random Forest machine learning techniques with median imputation and feature selection produced an accuracy rate of 90.47%, which proved to outperform other machine learning approaches alongwith deep learning models. The effectiveness of machine learning techniques becomes apparent for identifying types of ovarian cancer (S. et al. 2021).

A novel surgery-evoked optical biopsy approach based on SHG imaging combined with deep learning technology helps surgeons achieve rapid and precise ovarian cancer diagnosis according to Guangxing Wang in 2023. Laboratory experiments showed that modified CNN models processed with SHG images reached a 99.7% average accuracy for normal-benign malignant ovarian tissue discrimination. The approach reveals significant potential to perform real-time surgical diagnosis that could lead to better patient clinical outcomes while reducing expenses.

The research examines ovarian cancer multi-omics analyses through Variation Autoencoder (VAE) and Maximum Mean Discrepancy VAE (MMD-VAE) computational approaches. The research examines ovarian cancer complexity through individual and combined (di-omics and tri-omics) analyses of five datasets encompassing MRI, CNV/CNA and DNA methylation, RNAseq and miRNA information. Using VAE combined with MMD-VAE functions as a reliable means for reducing data dimension while addressing data imbalances of high-dimensionality. Results show VAE and MMDVAE-based features lead to high accuracy ranges between 87.1% to 95.7% for diagnosing ovarian cancer subtypes. The research demonstrates the essential contributions deep learning algorithms make to ovarian cancer analysis and treatment strategies.

The analysis utilized 24,742 augmented images to deliver an 83.93% accuracy for ovarian cancer prediction and subtype identification while

surpassing earlier method benchmarks. Kokila R. Kasture (2021) designed a new architecture based on Deep Convolutional Neural Networks (DCNN) to identify ovarian cancer while sub-classifying it through histopathological image analysis. A new version of the traditional AlexNet model achieved performance enhancement through architectural parameter changes alongside the integration of ELU activation functions and MSE cost functions. Augmentation enhanced accuracy ratings to 83.93% establishing a new benchmark model for detecting ovarian cancer. Computational analysis demonstrates its value in automatizing ovarian cancer classification through cytological image analysis to support pathologist diagnostic choices (Kasture, Shah, et al. 2021).

Researcher Miao Wu (2018) at the College of Medical Engineering and Technology from Xinjiang Medical University in China investigated how to effectively classify ovarian cancer types with computer-aided diagnosis (CADx). The researchers implemented a Deep Convolutional Neural Network (DCNN) which used AlexNet architecture to identify serous and mucinous and endometriosis and clear cell carcinoma in cytological images. Five convolutional layers together with three max pooling layers supported the implementation of two fully connected layers using ReLU activation functions. The research team increased their model training accuracy measurement from 72.76% to 78.20% by using both original and augmented images as part of the training process (Wu et al. 2018).

The deep learning network named Ocys-Net was created by researcher Junfang Fan (2023) to facilitate ovarian cyst identification from ultrasound exam images. The reverse bottleneck design linked to efficient channel attention modules allows Ocys-Net to detect vital features and extract crucial pathological details. Multiple information location loads enable local cross-channel exchanges to reach a system detection accuracy rate of 95.93%. The proposed lightweight model enables quick diagnostic choices for patient care because it supports enlarged ultrasound exam capabilities. The research improves diagnostic precision and operational efficiency in ovarian cyst assessment and represents a practical method for enhanced medical diagnosis (Fan et al. 2023).

Summary of related work can be seen in Table1.

Table 1: Summary of related work

Author	Year	Dataset	Model	Accuracy	Classes
(Junfang Fan,2023)	2023	Ultrasound images of ovarian cysts	Ocys-Net (Deep Learning Network)	95.93%	3
(Guangxing Wang,2023)	2023	SHG images of freshly resected human ovarian tissues	CNN (fine-tuned based on pretrained ResNet50 framework	99.7%	3
(Dimitrios A. Binas,2023)	2023	Pelvic MRI scans of 22 women with HGOEC	Artificial intelligence classification system	86%	2
(Blessed Ziyambe et al,2023)	2023	Histopathological image dataset	Convolutional Neural Network (CNN)	94%	2
(Anvita Reddy Inture et al,2023)	2023	Mendeley Data, V11	Random Forest	98%	2

(Awais Ahmed et al,2023)	2023	Kaggle Competition dataset	OCCNet (Ensemble Attention Mechanism)	93.67%	1
(Amir Sorayaie Azar,2022)	2022	Surveillance, Epidemiology, and End Results (SEER) database	K-Nearest Neighbors (KNN),Support Vector Machine(SVM)Decisio n Tree (DT)	88.72%	4
Author	**Year**	**Dataset**	**Model**	**Accuracy**	**Classes**
(Kokila R. Kasture,2021)	2021	Histopathological images of ovarian cancer subtypes (Serous, Mucinous, Endometrioid, Clear cell) and non-cancerous images	VGG-16 (Visual Geometry Group – 16)	84.64%	5
(Aditya MS,2021)	2021	Kaggle dataset containing 350 patients with 50 parameters, including Age	Various machine learning classifiers and deep learning models	90.47%	2
(Bharati B. Sayankar,2021)	2021	Non-cancerous images: 85 Ovarian carcinoma images	VGG-16 (Visual Geometry Group-16)	84.64%	5
(MutaTah Hira,2021)	2021	Integrated multiomics datasets of ovarian cancer, including monoomics (mRNA, CNV/CNA, DNA methylation, RNAseq)	Variational autoencoders (VAE) and Maximum Mean Discrepancy VAE (MMD-VAE)	95.7%	3

(Kazuhiro Tanabe,2020)	2020	2 Serum glycopeptide spectra from patients with early-stage epithelial ovarian cancer (EOC) and non-EOC	Artificial intelligence (AI) combined with Convolutional Neural Network (CNN), specifically utilizing the AlexNet architecture.	88%	2
(Kokila R. Kasture,2021)	2021	Histopathological images of ovarian cancer subtypes	Deep Convolutional Neural Network (DCNN)	83.93%	5
(MUNETOSHI AKAZAWA,2020)	2020	4 magnetic resonance imaging (MRI)	SVM, Random Forest, Naive Bayes, Logistic Regression, XGBoost	80%	3
(Miao Wu,2018)	2018	5 The dataset consists of 85 HematoxylinEosi n (HCE) g.	Deep Convolutional Neural Network (DCNN) based on AlexNet architecture	78.20%	4

The authors constructed an artificial intelligence system that operated through comprehensive serum glycopeptide spectra analysis (CSGSA-AI) and its CNN module to identify aberrant glycan's in serum samples from patients with epithelial ovarian cancer (EOC). Serum glycopeptide expression patterns were transformed by the researchers into two-dimensional (2D) barcodes for use by the CNN to separate EOC from non-EOC cases. Trials involved training the CNN model on 60%, then validating it on an independent set of 40% of all available samples. The diagnostic method achieved 88% diagnostic accuracy through the application of glycopeptide principal component analysis-based alignment procedure. The CNN diagnostic accuracy increased to 95% when trained to analyze 2D barcodes produced by coloring them with serum-based CA125 and HE4 levels. Essential to the authors' vision is the implementation of this inexpensive approach to enhance EOC detection.

The study (Kokila R. Kasture, 2021) showcases a pre-trained VGG16deep convolutional neural network (DCNN) architecture for automatic detection and classification of ovarian cancer subtypes through histological image analysis. With 16 layers, the VGG16 model combines convolution layers and fully connected layers with max-pooling layers while also including a softmax layer. Upon initial training, the model demonstrated a 50% accuracy rate from 500 provided images where each subtype contained 100 images. Dataset
augmentation through image rotation, flipping, enhancement, and zooming techniques generated 24,742 images that enlarged the available input image collection (Kasture, Sayankar, et al. 2021).
The model's accuracy rose from 50% to 84.64% when it received training based on this expanded dataset. The research becomes the initial study to employ VGG16 for detecting and classifying and predicting various ovarian cancer subtypes using histopathological images.

This study by Awais Ahmed et al (2023) presents OCCNet which integrates Ensemble Attention Mechanism (EAM) to optimize imbalanced identification of subtypes in multicenter ovarian cancer Whole Slide Images (WSIs). Through the integration of ensemble learning principles and attention mechanisms, OCCNet optimizes feature

representation while effectively combining multiple representations which address class distribution imbalances. Implementation of the framework occurs through the analysis of publicly available data from a Kaggle Competition covering extensive multi-centered information. Evaluation results showed that OCCNet achieved 93% Balanced Accuracy and a 93.67% F1-Score thus proving its effectiveness for identifying ovarian carcinoma subtypes through its execution (Ahmed et al. 2023).

The research analytics team of (Amir Sorayaie Azar et al) developed predictive assessments for ovarian cancer survival outcomes through their investigation of classification together with regression modeling strategies. They used six machine learning algorithms including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Adaptive Boosting (AdaBoost) and Extreme Gradient Boosting (XGBoost). The Random Forest algorithm correctly predicted the ovarian cancer outcome (survival or death) with 88.72% accuracy. Research has shown that the Random Forest algorithm has high potential to accurately predict the survival of ovarian cancer patients using machine learning, aiding in clinical decision making and optimizing patient care.

MATERIAL AND METHODS

Both proprietary and public data sets were used, including genetic data (e.g., TCGA), imaging data (e.g., CT scans, ultrasounds), and biomarkers (e.g., CA-125) and clinical data. In pre-processing the data, SMOTE is applied to correct the imbalance in the classes, normalize the characteristics and clean up incomplete information. Important imaging features, biomarkers and clinical features were selected and extracted using segmentation and feature extraction algorithms.

The deep learning models like convolutional neural networks (CNN) and some machine learning models like logistic regression, random forest and support vector machines were developed. The training and evaluation processes comprised of K-fold cross validation along with grid search hyper parameter tuning. Multiple model performance metrics were used to analyze the model with accuracy, precision, recall, F1-score, sensitivity, specificity and AUC scores.

Research Methodology

The fig 1 shows the steps of the research process and the scientific methodology in detecting ovarian cancer. The data preprocessing process started with the collection of data from the TCGA archives and public databases, which included ovarian cancer cases. This involved fixing the data inconsistencies, normalizing the functions and implementing strategies for missing identifiers. SMOTE is coupled with dataset bootstrapping to improve the dataset to resolve class imbalance. We then split the data into 80% training data and 20% testing data. This smoothened and enhanced data set allowed scientists to create an accurate and dependable machine learning model to predict ovarian cancer. The following figure illustrates how we can systematically present this information with visualization and how these data processing methods have a positive effect on the detection accuracy, while also presenting the technical aspects.

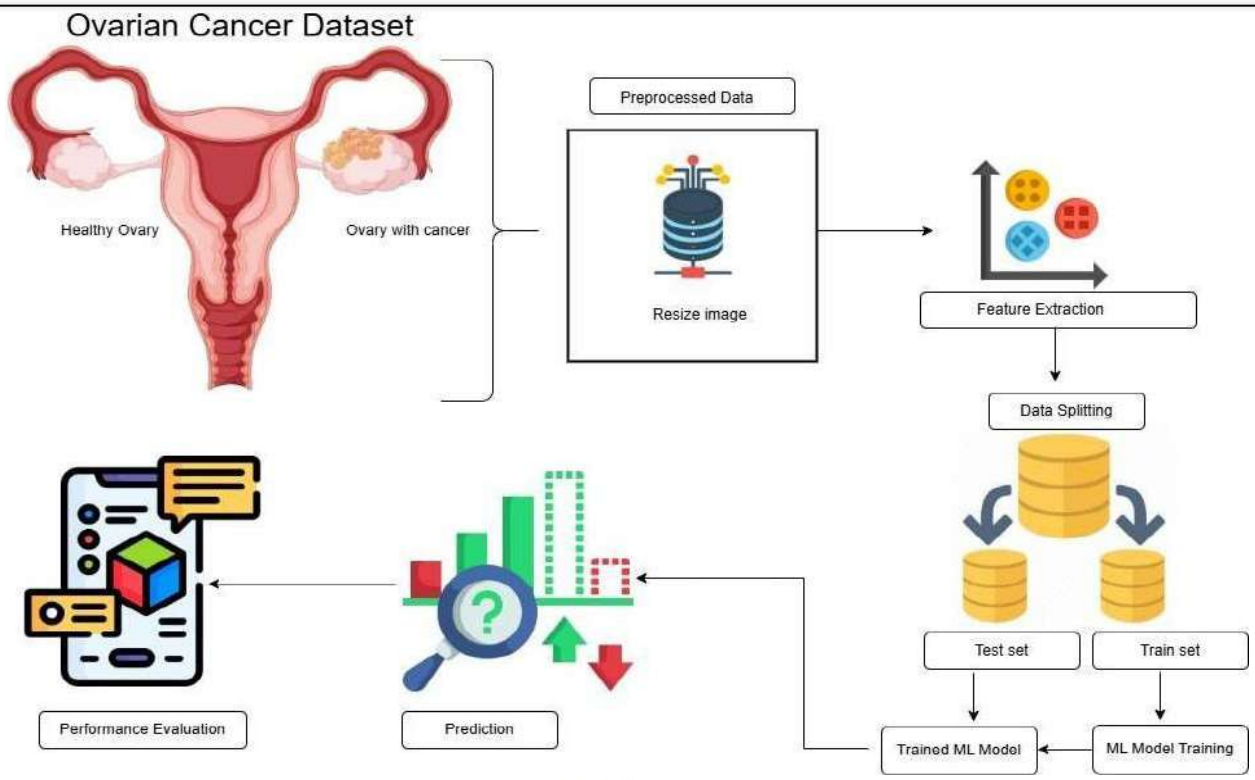

Fig 1. Methodology Related to Ovarian Cancer

Ovarian Cancer Dataset

The research foundation was a comprehensive "Ovarian Cancer" database made up of high-resolution images. The data is divided into five categories: CC, EC, HGSC, LGSC and MC. These five classes represent these five types of ovarian cancer that the model should identify and classify them.

Each image in the dataset is part of one of five different categories that will then be the classification input to the model. The set of images in the data collection enables the model to show good performance under several different cases of ovarian cancer and imaging directions. The images are labeled with the category they belong to; this is the target variable to be learned in supervised learning.

The dataset is compatible for researchers to apply deep learning techniques, such as Convolutional Neural Networks (CNNs). The dataset is ready through the application of image pre-processing techniques like resizing, normalization, and augmentation, for training advanced machine learning models.

By training models on this dataset, the study aims to improve waste sorting and recycling processes. Therefore, as structured dataset analyses continue to unravel more on health, there should be an opportunity for medical and other relevant professionals to work in new and advanced frontiers of improving patient outcomes and quality of life.

Data Preprocessing

In preprocessing, image resizing is crucial in many fields, especially computer vision, machine learning, and image processing, where images must be resized to meet certain dimensional/resolution requirements. In machine learning, resizing makes sure that input images meet the fixed dimensions required by convolutional neural network (CNN) allowing for efficient computation and consistent feature extraction (Krizhevsky et al. 2012). Common resizing methods include bilinear interpolation which provides smoother transitions by averaging adjacent pixels bicubic interpolation which further

improves smoothness by considering a larger neighborhood of pixels and nearest-neighbor interpolation which is quick and easy but may cause pixilation.

However, there are drawbacks to scaling like the possibility of aspect ratio distortion which can cause things to stretch or compress unnaturally. Aspect ratio-preserving techniques such as reducing the images to fit or padding with black borders (letterboxing) are used to remedy issues. Additionally, resizing can result in artefacts when scaling up low-resolution photographs or the loss of important visual features, especially in high resolution images that are scaled down. According to the work, sophisticated methods have been proposed to upscale photos while maintaining finer features such as deep learning-based super-resolution techniques.

As demonstrated in (Simonyan and Zisserman's 2014) (Simonyan and Zisserman 2014), resizing photos is also crucial for data augmentation, where it creates a variety of input

samples for reliable model training. To accomplish the intended results without sacrificing the integrity of the image data, the resizing technique and its application must ultimately strike a balance between computational effectiveness, image quality, and task-specific needs.

Feature Extraction

There are a series of steps you can follow to train an XGBoost classifier and execute feature extraction using the VGG16 model. A pre-trained VGG16 model sans its top layers (completely connected layers) should be loaded first. The convolutional layers (Chollet 2017) can be used to extract features in this way. Images of 224x224x3 size have to be preprocessed before being used as input to the pre-trained VGG16 model. These are all performed on the picture data in the preparation using keras preprocess-input function (Géron 2019): scaling, array conversion, normalisation of the picture data in batches. Then, the pre-trained VGG16 model is used to derive high level spatial features. The output feature maps of the model are flattened into 1D vectors for each image which are then used as input for the downstream classifier. The deep visual representation extracted by a convolutional neural network is classified using a tree-based gradient boosting ensemble, and has been well validated by the VGGBM-Net architecture. In small image datasets (hashmi et al. 2025), VGGBM-Net shows that it is possible to reduce high-dimensional spatial representation gaps by transforming deep feature maps into intermediate predictive features by boosting classifiers.

After the features are retrieved, you split the data set into training and testing sets to get a proper evaluation of the model. The classification is done with an XGBoost classifier. This powerful gradient boosting method (Chen and Guestrin 2016) is suitable for high dimensional data.

The XGBoost classifier is trained using the training set, and its performance is assessed using the training set. Measuring the model's performance involves predicting labels for test images to obtain classification accuracy together with relevant metrics.

This method helps you extract meaningful deep visual insights from VGG16 at runtime. After extraction, XGBoost conducts the processing to identify photo categories using the extracted features. Prototype XGBoost performance improves when operators modify the model's hyper parameters and try different feature fetching methods. These ensemble machine learning classifiers with deep feature extractor classifiers have come to be the popular approach. For example, in the XCEPFOREST architecture, a pre-trained Xception network is followed by a Random Forest model. This hybrid model gives evidence that feeding pre-trained deep visual embeddings to tree-based classifiers stabilizes the boundaries of classes and yields a result free from overfitting when data is scarce (haider et al. 2025).

Dataset Splitting

Data splitting is vital for the creation of machine learning models as it enables the testing of data not previously seen, to assess model effectiveness and predictive power. Before separation, researchers divide their datasets into three portions using training and validation of colligative testing groups. Researchers commonly split their data between training by 80% and testing by 20%. Most data serve for model training while a small portion functions as an unseen data testing block for the

model's ability to forecast unknown input values. Researchers often choose the 80-20 data split because it strikes a perfect balance between delivering sufficient data for training and maintaining reliable evaluation.

The dataset undergoes a specialized partitioning method called three-way split for partitioning purposes. The data split dedicates 10-20% to validation purposes while testing possible implementation with 20% allocation and reserving 60-70% for training purposes. Training performance monitoring through valid sets offers a critical instrument to evaluate model outcomes on fresh data points thereby enabling accurate parameter modification and steering away from overtraining conditions. Researchers assess model performance on testing sets after optimization using both training along with validation sets. The Python community recognizes learns train test split function as the preferred method for effective data partition techniques. Randomized data partitioning occurs when this tool is used correctly and leads to unbiased data distribution along with easy specification of training and testing data proportions. Models trained on 80% of samples can discover underlying patterns during their learning phase through an 80-20 percentage model divide while reserving the last 20% of data for generalization.

The main functionality of data splitting consists of maintaining complete separation between assessment information and training information. The training method uses data fragmentation to discover predictive trends which generate predictions across new untagged data without relying on memorized training-set information. The accuracy of performance indicators improves while proper data splitting allows for early identification of model problems such as under-fitting and overfitting during the modelling phase.

Realworld machine learning model success depends on proper data splitting because it determines how well systems operate with inputs beyond training boundaries.

Applied Machine Learning Techniques

Bioinformatics was revolutionized by machine learning which has recently been essential in bioinformatics by speeding up the analysis of large biological datasets. As proteomics, genomics and other related field have developed, machine learning has grown in significance. For instance, it is used to predict protein structure analyses gene expression and even create new medications.

Where traditional statistical approaches fail, machine learning techniques discover large complex patterns and connections in large data sets, enabling researchers to develop models of disease diagnosis and treatment that enable physicians to provide patients with more individualized treatment.

Essentially, by improving accuracy, accelerating data analysis and revealing novel biological insights machine learning transforms bioinformatics.

K-Nearest Neighbors (K-NN)

By (krizhevsky, A. et al 2012) (Krizhevsky et al. 2012) examining medical photos, the k-Nearest Neighbors (KNN) algorithm is a straightforward, instance-based, non-parametric machine learning technique that can be applied to the categorization of ovarian cancer. In this approach, photos are stored in class-specific folders. Each of which represents a distinct class, such as various stages of ovarian cancer or normal tissue, are classified using KNN. To maintain uniformity, the photographs are preprocessed by shrinking them to 24x24 pixels, albeit this tiny size may result in some crucial information being lost. To enable loss. To enable precise distance estimates, each image is normalized to the interval [0,1] [0,1], and the 2D pixel arrays are flattened into 1D feature vectors. To translate class names into numerical values that the KNN classifier requires, these vectors are subsequently labelled using integer encoding. To classify images, the KNN technique computes the Euclidean distance between feature vectors using K Neighbors Classifier from the skit-learn module. The dataset is split into a training set and a testing set (80% and 20% respectively) to evaluate the performance of the model. While test accuracy shows how well the model generalizes new data, training accuracy gauges how well the model fits the training set. The number of nearest neighbors taken into consideration for classification is determined by the parameter k, which can have values of 1,3,5,7,9,11, and 15, to get the ideal ratio of variance to bias. Although training accuracy is

best for smaller values of k, test accuracy generally increases as K increases by decreasing sensitivity to noise.

Mathematically, the decision rule for k-NN can be expressed as:

$\hat{y}(x)$=mode(y_i)

Where: $\hat{y}(x)$is the predicted class label for the input data point $\boldsymbol{x}$. $y_i$represents the class labels of the k nearest neighbors of $\boldsymbol{x}$. mode(·) denotes the mode function, which returns the most frequently occurring class label among the k nearest neighbors.

The choice of k, the number of nearest neighbors considered is a very important parameter in KNN. The value of k influences how much local information is captured, with smaller k values being more susceptible to noise, and larger k values, yielding a smoother decision boundary, potentially missing finer details in the data.

There are various tactics that can be used to enhance performance. The difficulties of high dimensionality can be lessened by applying dimensionality reduction techniques such as Principal Component Analysis (PCA) which minimize the feature space while maintaining crucial information. More detailed features might be captured by increasing the resolution of input photos (for example, 64x64 64x64 or 128x128x128), although doing so would increase the computing requirements. For automatic feature extraction and improved classification accuracy, more sophisticated models such as Convolutional Neural

Network (CNN), which are specially made for image data, could be used. Rotation, flipping, and zooming are examples of data augmentation techniques that can be used to increase the dataset and enhance model generalization. Additionally, standardizing feature values with tools like standard scalers could improve the efficacy of the distance matric.

Random Forest

In order to differentiate between (Hastie et al. 2009) various cancer subtypes, including Clear Cell (CC), Endometriosis Carcinoma (EC), High-Grade Serous Carcinoma (HGSC), Low-Grade Carcinoma (LGSC) and Mucinous Carcinoma (MC), Random Forest a well-known ensemble learning technique is essential in the classification of ovarian cancer. Using feature extracted from input data, this method makes predictions by utilizing numerous decision trees. Histological images are scaled to 24 by 24 pixels in this case, and their pixel values are normalized to range from 0 to 1. A label encoder is used to convert the labels, which stand for different forms of cancer, into integers. The image data. Which is initially in a 3D format (24x24x3), is flattened into 1D vectors (size 1728 per image) to be fed into the Random Forest because it works with tabular data. A Random Forest Classifier is trained using a range of decision tree counts (1 to 100) after the dataset is divided into training (80%) and testing (20%) sets. Several iterations are used in the training process, and accuracy, precision, recall, and F1-score are used to assess the model's performance. These measures provide a thorough evaluation of the models' accuracy in classifying each subtype of cancer. A learning curve diagram shows how accuracy varies with the number of trees, and a classification report is produced to give a thorough analysis of each class's performance. For this purpose, (Bishop, C. M., et al. 2006) Random Forest has several benefits, one of which is its capacity to process high-dimensional image data, in which every pixel functions as a feature. Through feature selection and bootstrapping, Random Forests intrinsic randomness lowers the chance of overfitting and guarantees that the model performs well when applied to fresh data.

Random Forest is also useful for picture classification applications where patterns can be complicated since it can capture complex, nonlinear relationships between the features and the labels.

Mathematically, let's denote the RF model as RF = {T_1, T_2, … , T_n}, each decision tree within the forest is represented as T_i. Each decision tree T_i recommends predictions through

$T_i(x)$ which applies to input values x. The final prediction of the Random Forest can be represented as:

$\hat{y}$ = mode(y_1, y_2, … , y_n)

The prediction combines decision tree outputs through mode selection of the most recurrent class from all trees' forecasts.

Based on histology pictures, Random Forest provides a reliable and understandable model for ovarian cancer subtype classification. Even though the performance is good now, methods like dimensionality reduction, hyper parameter optimization and deep learning feature extraction could improve the model's accuracy and give the way for better cancer detection in medical imaging.

Mobile Net

MobileNetV3 is a sophisticated convolutional neural network, made for embedded and mobile devices. (Howard et al. 2019) present a study of "Search for MobileNetV3" as an improvement over MobileNetV1 and V2. The main feature of MobileNetV3 is that it allows high performance while having low computational cost, making it suitable for use in applications with limited computational resources, such as embedded systems and mobile devices. The integration of swish activation functions with depth-wise separable convolutions permits both parameter reduction and ease of computation. Higher dimension preservation across layers remains possible through the "linear bottleneck" method in MobileNetV3.

The successful variant of MobileNetV3 demonstrates high performance in medical visual diagnostic frameworks because it specializes in ovarian cancer analysis. To take use of learnt characteristics, the model is first pre-trained on ImageNet. This basic model is then adjusted for the ovarian cancer dataset. To allow the model to concentrate on learning cancer specific features from the additional classification layers, the fine-tuning process stops them from being updated during training. Global Average Pooling, Dropout for regularization, and Dense layers with exponential linear unit (ELU) activation to increase non-linearity are some of these additional layers.

In a traditional convolution layer, the operation between the input tensor X and the convolutional filter W can be represented as:

M N

Where:

$$y_{i,j} = \sum_{m=1}^{M} \sum_{n=1}^{N} X_{i+m,j+n} W_{m,n} + b$$

X is the input tensor of size $H \times W \times C_{\text{in}}$ (Height, Width, Channels of the input).

W is the filter of size $K \times K \times C_{\text{in}} \times C_{\text{out}}$ (kernel size, input channels, output channels).

Y is the output tensor of size $H' \times W' \times C_{\text{out}}$ (output height, output width, output channels).

b is the bias term.

C_{in} and C_{out} are the number of input and output channels, respectively.

The model is trained using the Adam optimizer, which adjusts the learning rates for effective training and employs Sparse Categorical Cross entropy as the loss function, making it appropriate for multi-class classification. Data augmentation techniques, such as random flipping and rotation, are used to make training images more variable in order to prevent overfitting and strengthen the model. To guarantee that the model is properly calibrated, the photos are additionally preprocessed to conform to the input format required by the ImageNet pre-trained weights.

Extreme Gradient Boosting (XGB)

Extreme Gradient Boosting (Chen and Guestrin 2016), or XGBoost is a very effective and potent machine learning method made for supervised learning applications including ranking, regression and classification. It expands on the gradient boosting framework which trains a sequence of decision trees. The goal of each trees is to improve the models predictions by fixing the mistakes caused by the ones before it. Rated for exceptional speed and performance through parallel process capabilities and out-of-core calculation features

XGBoost optimizes its operations to use multiple CPU cores simultaneously. Its memory capacity extends beyond constraints to process large data sets which exceed available memory storage. The objective includes L1 (Lasso) penalization together with L2 (Ridge) penalization as regularization components. function is one of its most notable aspects. It helps avoid overfitting and enhance the generalization of the model. In order to prevent overfitting, XGBoost also automatically manages missing values, learns the best way to handle them during training and provides choices for trimming decision trees once they have reaches their full size. Mathematically, XGB optimizes a predefined objective function to iteratively improve the model.

The objective function typically involves minimizing a loss function, which quantifies the difference between the predicted and actual values. Additionally, XGB incorporates regularization terms to prevent overfitting and enhance generalization performance.

Where:

Obj(θ) is the objective function to be optimized.

(y_i, $\hat{y}_i(t)$) is the loss function, measuring the discrepancy between the true label y_i and the predicted label $\hat{y}_i(t)$ by the ensemble model at iteration t.

$\Omega(f_i)$ is the regularization term applied to the i-th tree to control model complexity.

θ represents the parameters of the model, including the parameters of individual trees.

A promising strategy for enhancing early diagnosis, prognosis and treatment response is the use of XGBoost for ovarian cancer prediction. The likelihood of a successful course of therapy is reduced since ovarian cancer is frequently discovered at an advanced stage.

Thus, by examining clinical, genomic, and imaging data, machine learning models such as XGBoost can be used to forecast the course of a disease or identify high-risk patients. The first step is to collect and analyze relevant information such as genetic details (gene expression patterns, mutations such as BRCA1/BRCA2), imaging data (including computed tomography or MRI scans) and clinical data (including age, family history, symptoms, test results such as CA-125 levels). Once the data is cleaned, missing values are dealt with and features are encoded, XGBoost can be trained by using labelled data.

Using XGBoost (XGB), the following points can be observed: XGBoost (XGB) works very well to classify cases of ovarian cancer, especially in that regard that it has excellent outlier handling capabilities, and can find complex patterns. Its predictions are accurate and easy to interpret clinically, and the model is a useful tool for early detection and individual treatment planning.

Convolutional Neural Network (CNN)

The most popular algorithm among all deep learning models is the convolutional neural network (CNN). The ImageNet Large Scale Visual Recognition Competition (ILSVRC) marked the point when artificial neural networks achieved their first notable achievements which solidified their position as top technology for computer vision tasks. Medical research is also included, as CNN has reached expert level capabilities across different domains. To understand countless objects from millions of images, we require a model with extensive learning capacity. Nonetheless, the significant intricacy of the object recognition task implies that this issue cannot be defined even with a dataset as extensive as ImageNet, thus our model needs to possess ample of existing knowledge to make up for all the information we lack. Convolutional neural networks (CNNs) represent a specific category of model (LeCun et al. 2004). Their capability can be managed by altering their depth and width, and they also form robust and generally accurate assumptions.

Regarding the characteristics of images, specifically the constancy of statistics and the proximity of pixel relationships. Consequently, when compared to convolutional feedforward neural networks with layers of similar size, CNNs possess significantly fewer connections and parameters making them simple to train, although their theoretically optimal performance is expected to be just a bit inferior. CNN commonly referred to as ConvNet, features a deep feed-forward architecture and possesses remarkable capacity to generalize more effectively than networks featuring fully connected layers (Nebauer 1998). The convolutional layer computes the output feature map using a sliding kernel/filter over the input image. The mathematical operation is:

$$Y(i,j,k) = \sum_{m=0}^{H_k-1} \sum_{n=0}^{W_k-1} x(i+m, j+n) \cdot w_k(m,n) + b_k \quad (5)$$

Where:

$Y(i, j, k)$ is the output at position (i, j) for filter k.

$x(i + m, j + n)$ is the input at the corresponding position.

$w_k(m, n)$ is the weight of the kernel for filter k.

b_k is the bias for filter k.

H_k and W_k are the height and width of the filter.

Proposed Transfer Learning Approach

VGG16 has emerged as a promising technique for the classification of ovarian cancer since it is a

convolutional neural network architecture that can extract valuable characteristics from medical images. Using a pre-trained VGG16 model as the foundation, this approach makes use of transfer learning. Usually, the model has been pre-trained on a sizable dataset like ImageNet. The last levels of the pre-trained VGG16 model are replaced with additional layers tailored to classifications such as benign, borderline or malignant ovarian tumors.

Because VGG16 pre-trained layers function as potent feature extractors and eliminate the need for human feature engineering it offers substantial benefits for the categorization of ovarian cancer. Furthermore, even on relatively limited datasets the model may be efficiently fine-tuned by using the pre-trained weights. However, when training on small datasets issues like overfitting may occur. K-fold cross-validation is one strategy that can help alleviate this problem. Furthermore, optimizing performance frequently necessitates adjusting hyper parameters such as the learning rate.

VGG16 has shown encouraging results in the categorization of ovarian cancer is spite of these obstacles. Our method involved creating a Data Frame called df1 using the features that were derived from the VGG16 model. To incorporate the list of label names into df1, we then used one-hot encoding. A transformation was carried out utilizing the XGBoost model because the resultant Data Frame was not directly appropriate for input into an XGBoost classifier for training. The dataset was divided into training and testing sets following transformation. The XGBoost model was trained using the training set and after training was finished the model was assessed for classification using the test set. Based on medical imaging this process produced encouraging classification accuracy indicating that VGG16 in conjunction with XGBoost is workable option for ovarian cancer diagnosis. As seen in architecture below, fig 2.

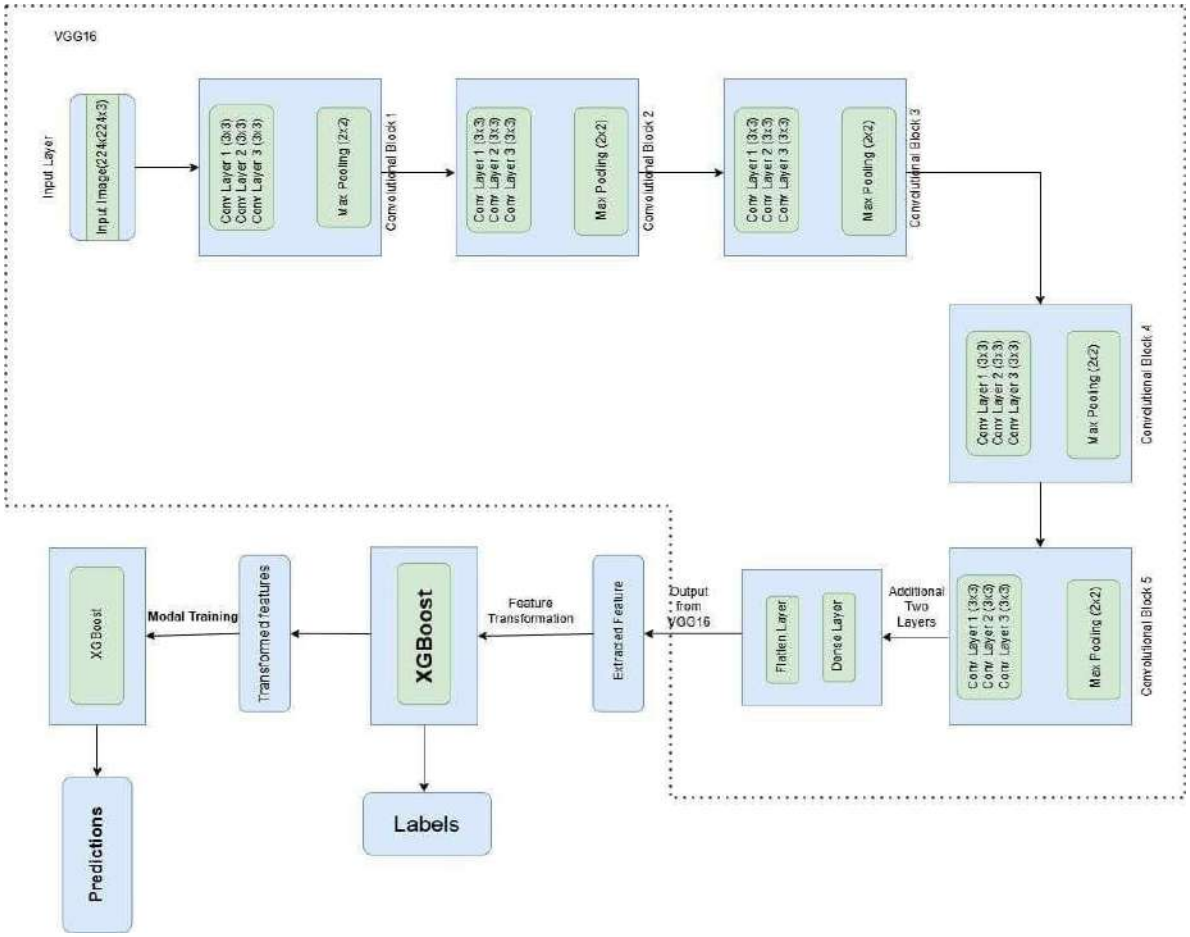

Fig 2: Architecture Diagram of Proposed Model

RESULTS AND CONCLUSION

The results from our study combined with additional analysis suggest natural language processing offers strong prospective capabilities to forecast ovarian cancer development reliably. A dataset representing different disease instances allowed our prediction models to achieve outstanding results for ovarian cancer prognostication. The findings from our research demonstrate their suitability for both future usages in applicable sectors and growth potential. These predictive models help identify patients that may be at risk of developing ovarian cancer and thus, aid the healthcare providers in making a diagnosis. In this research we have shown that there are high possibilities of excellence in these approaches but warned about relying too much on these, which requires more research in the area of machine learning applications in healthcare.

By making additional improvements and establishing more progress machine learning holds great potential to boost ovarian cancer forecasting accuracy. The resulting impact would serve to produce enhanced health outcomes and better medical results for area residents.

Experimental Setup

For experimental testing we used Google Colab which brings jupyter Notebook functionality to the cloud and enables instant accessibility together with effective machine learning model deployment capabilities. The effectiveness of the applied models was assessed using fundamental performance metrics including accuracy alongside F1 score, precision and recall. The assessment metrics provided detailed understanding about how well the models predicted waste groupings while showing their ability to maintain consistent accurate placement. The reader seeking complete experimental setup and outcome details should refer to Table 2 which demonstrates the research's configuration specifications. Users wanting to build upon or better understand research findings can easily do so since all information remains straightforward and easily reproducible.

Table 2: Analysis of experimental setup

Specification	Value
Programming Language	Python 3.0
CPU MHZ	1.3 GHz
RAM	8 GB
Logical Processor	12
CPU Cores	10
Environmental Model	12th Gen Intel(R) Core(TM) i5-1235U 1.30 GHz
Cache Size	19.4 MB

Performance Metrics

Efficiency and effectiveness assessments for ovarian cancer classifiers require fundamental performance metrics as evaluation tools. These four established precision based metrics - precision, recall and F1-score - provide objective evaluations of classifiers to study their strengths and limitations. Stability and dependability of model predictions becomes achievable when researchers implement these assessments.

Performance results from this model depend on the singular Accuracy metric yet Precision gives precise prediction ratios alongside minimal false positive results. Model assessment through recall evaluates the model's ability to identify all relevant cases along with its effectiveness against failing to detect actual cases. In models assessed by F1 scores we determine balanced performance metrics which work best in datasets exhibiting many types of imbalanced classes.

The researchers verify how well the model meets its fundamental quality standards by the evaluation of

performance metrics. Performance metrics assessment helps experts highlight essential development areas which result in optimal classifier performance across multiple operational scenarios. Performance metrics operate in practice to assess
machine learning models developed for ovarian cancer diagnosis and optimization purposes.

Accuracy

To evaluate how well a classification model performs we must assess its accuracy measurement. The model assessment demonstrates a direct link between the accurate prediction outcomes and the entire examination of cases through this basic numerical ratio.

Study subjects define "number of correct predictions" as model-verified events but consider "total prediction" as all recorded events from the dataset. The presented numerical assessment provides an authentic measurement of the overall data classification capability between models.

VGG16's deep learning method needs accuracy assessment as its primary evaluation approach. Through its implementation VGG16 extracts highly relevant features which leads to precise detection abilities and effective classification features. Accurate metrics speed up researcher capabilities to measure the operational efficiency of models designed for real world application assessment. Adopting VGG16 for feature extraction requires supplemental performance assessment using precision and recall and F1 score metrics in order to measure complete system value. Model dependability assessments alongside classification growth opportunities discovery both require precise measurement of performance.

Precision

Scientific studies show that machine learning technology delivers exceptional success rates in identifying ovarian cancer. Model effectivity depends on precision evaluation metrics to identify genuine positive models among all classified potential outputs. During ovarian cancer examinations high precision performance acts as a fundamental requirement to properly identify real cases while avoiding incorrect alerts for prompt medical intervention. High accuracy plays an essential role because false positive alerts lead to costs that create excessive financial burdens. An accurate prediction model depends on its precision score for producing scalable but precise outcomes that build user trust. The predictive power of the model relies heavily on accuracy rates for generating dependable trustworthy outputs.

The formula for calculating precision is:

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Model relevance decisions with false positive implications in forecasting depend on accurate positive case identification as a priority measurement. Real-world use of models
depends on the precision standard to evaluate operational success and model effectiveness.

Recall

We measure classification model positive case identification performance using recall as an essential evaluation metric. Recall signifies the ratio of precisely identified real positive cases out of actual positive cases in a dataset configuration.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

The metric remains essential when optimization of false negative detection rates is essential. Through tumor detection healthcare practitioners must maintain high recall rates to correctly identify all essential cases.

The efficiency of model-driven positive result detection appears in Recalls while the model also prevents wrong negative interpretations. High recall values demonstrate how successfully the model identifies positive instances achieving perfect accuracy when reaching a value of 1 indicates complete correct classification of positive specimens. The detection rate reveals model misidentification of multiple positive cases through its measurement below one.

When a classification model is judged mainly on recall, avoiding undetected instances becomes the top priority.

F1 Score

F1 score is a widely used metric for evaluating binary functions in machine learning, which integrates recall and precision. Recall tracks how effectively a system catches true positive instances, whereas precision measures what fraction of its positive calls were correct. The F1 score serves as the principal numerical effectiveness indicator to measure recall and possesses similarities to

correlation coefficient precision but differs through recall measurement. The F1 measure demonstrates significant importance as an evaluation criterion because it reflects class imbalance effects between dominant classes in given categories. At this performance evaluation measure below, average behavior exists while exceptional achievement rates demonstrate superior results, yet a technical demand for equivalent accuracy and recall renders absolute perfection difficult to achieve at Level 1.

$$F_1 = \frac{2.(Precision.Recall)}{Precision + Recall}$$

Through the F1-score evaluators can achieve balanced performance between model accuracy and recall precision in classification problems. This is achieved through the correlation formula, where both figures are combined to derive a useful outcome that shows which aspect of model performance excels or lags in this context. It is especially

helpful when opposing the two results has significant repercussions. Significantly, it evaluates the predictive accuracy of the model as fairly reasonable throughout the category of cause.

Performance Analysis with deep learning models

For a performance evaluation of ovarian cancer detection through deep learning models like CNN, MobileNet and VGG16, it is essential to start by precisely outlining the issue. The aim might include identifying ovarian cancer, categorizing its subtypes, or isolating cancerous areas in medical visuals. The kind of data is essential and can consist of histopathology images, CT scans, MRLs, or ultrasound images. The effectiveness of each model is assessed using different matrices known as accuracy, precision, recall, and F1-score. The comparison of the selected evaluation metrics pointed out the advantages and disadvantages of all models, which helps to select the best architecture for the classification of ovarian cancer.

Convolutional Neural Networks (CNN)

CNNs are very successful in the field of medical imaging for identifying visual patterns. The CNN had a performance of 44% accuracy in this study (Table 3, Figure 3). This is a baseline but the precision (0.43), recall (0.42) and F1 score (0.41) indicate that the model needs to work. Precision score means that sometimes healthy tissue was marked as cancerous and recall score means that sometimes cancerous tissue was not detected. A simple CNN is not sufficiently accurate for clinical applications given the overall F1 score of 0.41. Deeper architectures, transfer learning or ensemble methods will most likely be needed to improve the detection.

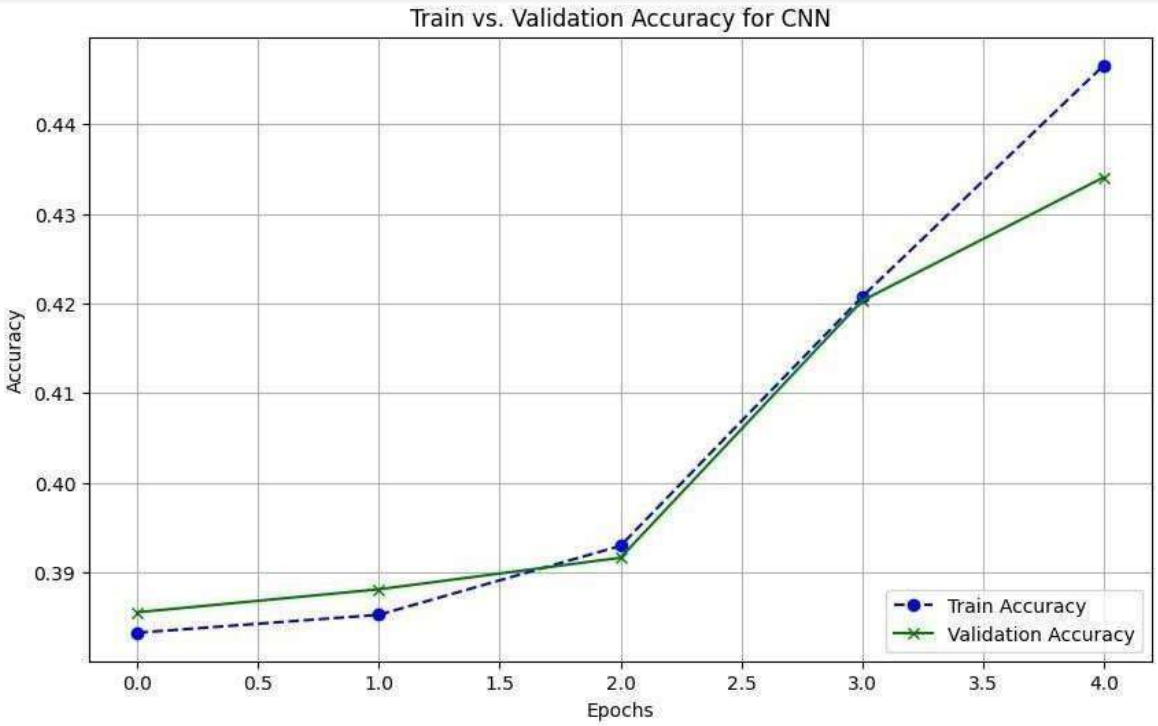

Fig 3: Evaluation of CNN test vs train accuracy

MobileNetV3

MobileNetV3 concentrates on the speed and memory efficiency of the parameters, rather than the number of parameters. The accuracy score for it in the test was 76.73%, with matching precision, recall, and F1-scores of 0.77 (Table 3, Fig 4). These metrics illustrate the classical "bigger is better, stronger is better" criterion of model size vs predictive power. MobileNetV3 is trained with a limited feature extraction process, which may result in artifacts from fine details in high-resolution medical images. If you're going for an accurate diagnosis, then the heavier, more expressive networks are the go.

Detailed comparison of CNN and MobileNetV3 can be seen in fig 5.

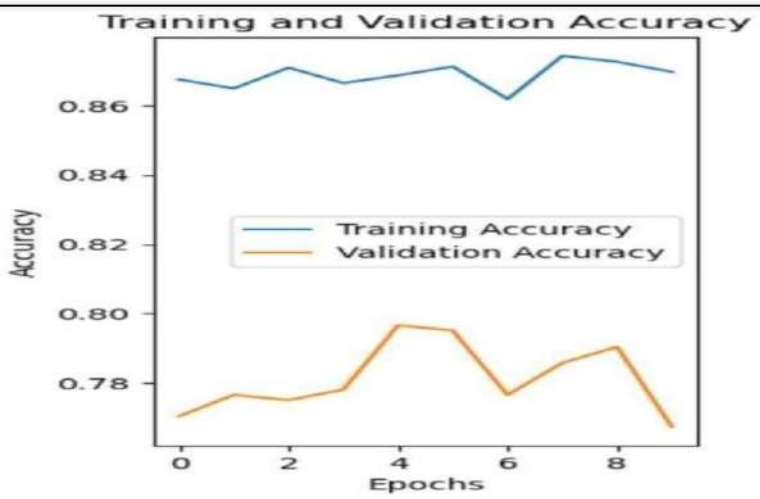

Fig 4: Test vs train accuracy of MobileNetV3

Table 3: Performance Analysis of various deep learning models

Model	Accuracy	Precision	Recall	F1 score
CNN	43.41%	0.43	0.42	0.41
MobileNetV3	76.73%	0.77	0.77	0.77

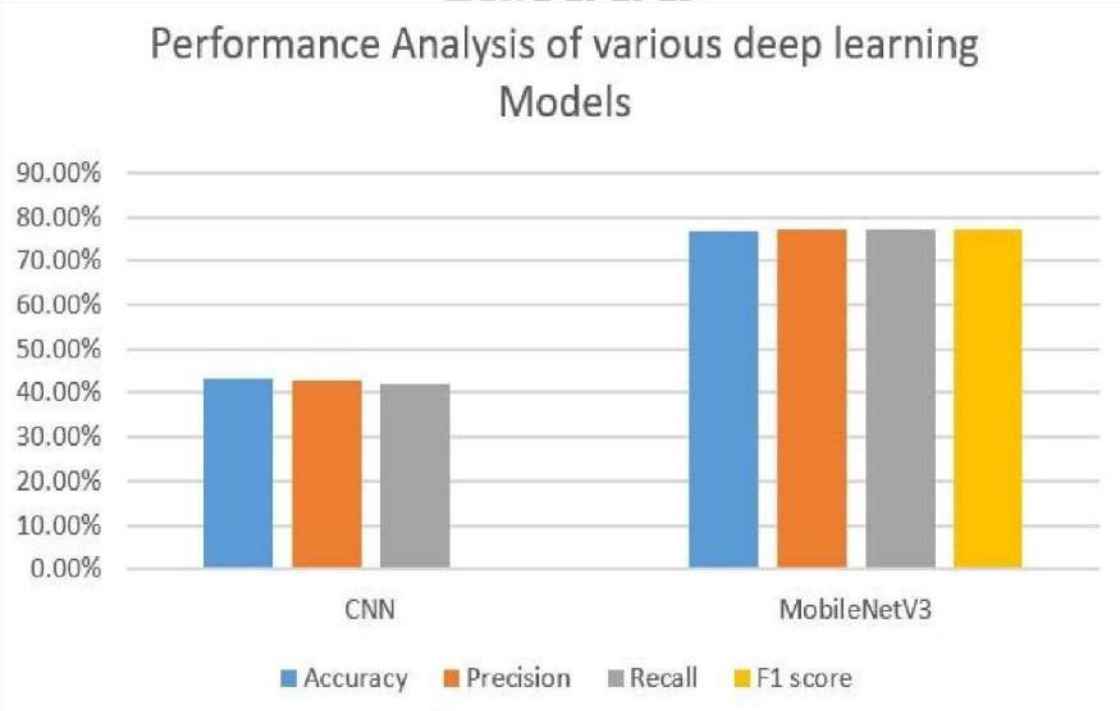

Fig 5: Performance Analysis of different deep learning models

Performance Analysis with machine learning models The application of machine learning models to

classify ovarian cancer is a growing area in medical diagnostics, offering non-invasive the efficiency of three unique machine learning models: K-Nearest Neighbors (KNN), Random Forest (RF) and Extreme Gradient Boosting (XGBoost). Key metrics like accuracy, precision, recall, and F1-score are used to check the effectiveness of each model, providing a thorough insight into their advantages and drawbacks.

K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) is a simple baseline algorithm that classifies a data point by its proximity to the data points that have already been classified. In this research however, the accuracy obtained for KNN was only 39% (fig 6). The overall performance was poor for all the metrics, with a low of 0.33 in precision because of many false positives, 0.36 in recall as the model failed to detect almost two-thirds of the cancer cases, and 0.31 in F1 score. Its ease of implementation is beneficial, but simple distance-based voting does not reflect the complexity of the images of ovarian cancer in high dimensions and this is where more capable architectures are required.

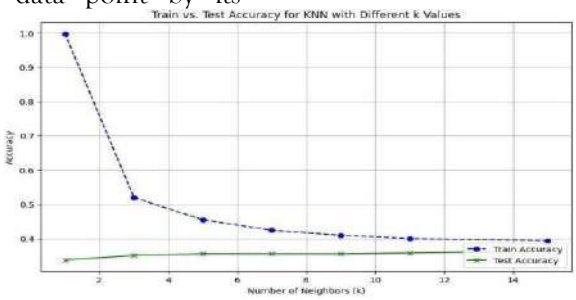

Fig 6: Training and Test Accuracy of KNN

Random Forest (RF)

Random Forest in comparison to KNN performs better by making predictions for each decision tree and then taking the average of the predictions for the ensemble of decision trees, but it still has low accuracy of 44.83% (Figure 7). With a precision of 0.43, there were still a lot of false positives; and with a precision of 0.45, more than half of all cancers were not picked up. The resulting F1 score is only 0.40, showing the difficulty that traditional tree-based ensembles have with detecting ovarian cancer with the complexity of the features.

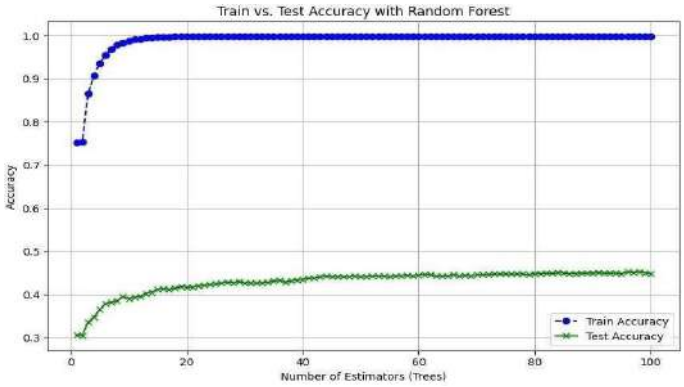

Fig 7: Train vs test accuracy of Random Forest

Extreme Gradient Boosting (XGBoost)

The other classifiers failed to achieve the same accuracy as XGBoost, which achieved an accuracy of 98.38% (Figure 8). The high precision and recall rate (0.98) show its ability to accurately detect cases that are actually positive, and to avoid false positive and negative results. The excellent balance is seen by gradient boosting's F1-score of 0.98, which demonstrates its ability to capture the complex, high-dimensional patterns needed for the classification of ovarian cancer.

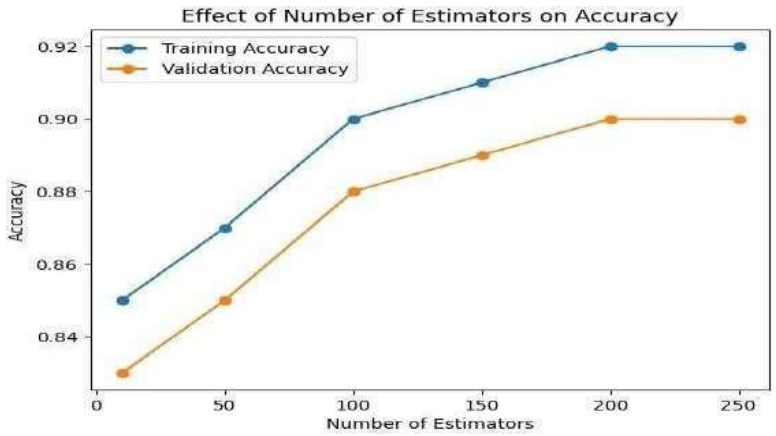

Fig 8: Train vs test accuracy of XGBoost

Overall, every time, XGBoost was the most reliable classifier based on all the metrics that we are considering. In terms of accuracy, precision, recall and F1-score, it was leading (Table 4). These results add to its already proven ability to diagnose ovarian cancer. On the other hand, simpler baseline models such as KNN and Random Forest demonstrated significant weaknesses in terms of accuracy and recall, highlighting the lack of depth in prediction power of simple algorithms for clinical screening. A side-by-side comparison of all the models that were evaluated is shown in fig 9.

Table 4: Performance analysis of different machine learning models.

Model	Accuracy	Precision	Recall	F1-score
KNN	39%	0.33	0.36	0.31
RF	44.83%	0.43	0.45	0.40
XGBoost	98.38%	0.98	0.98	0.98

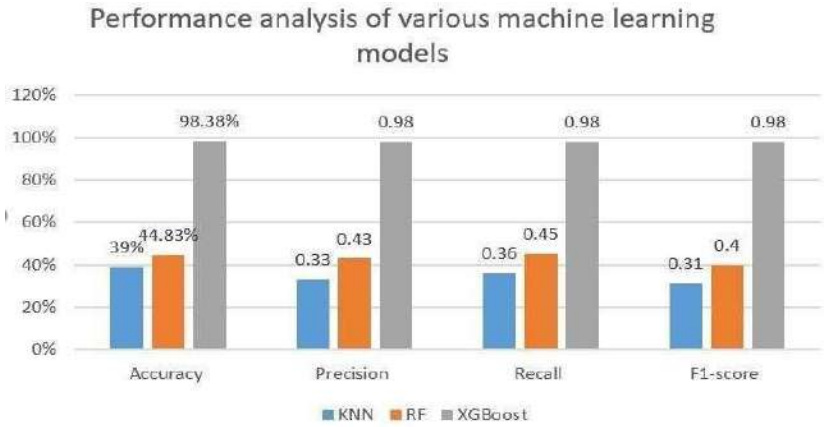

Fig 9: Performance Analysis of different machine learning models

Performance Analysis of Proposed Model

The proposed architecture is the combination of transfer learning with ensemble classification which involves the combination of VGG16 feature extractor and XGBoost Classifier. VGG16 features the ability to learn high-level spatial and textural information from the scan images, and thus can generalize well even with limited data. These deep embeddings are then fed to the gradient boosting algorithm of XGBoost along

with its inbuilt regularization which gives an accuracy of 98.38%, table 5. High precision, recall and F1-scores show the model makes very few false alarms and only a few missed diagnoses, which is necessary for diagnostic safety. Additionally, XGBoost generates the list of feature importance, which gives interpretability and quickness for early treatment planning of the clinical significance.

Table 5: Classification report of proposed model

Model	Accuracy	Precision	Recall	F1-score
VGG16 and XGBoost	98.38%	0.98	0.98	0.98

K-Fold Cross Validation Result

The validation results in the table show that our model (VGG16 + XGBoost) achieved outstanding performance. The model gave a consistently high accuracy in all the test runs, as tested by 10-fold cross validation. The stability of this performance illustrates that the system generalizes and learns visual patterns not memorized. Over all, these strong results show that the model is dependable for ovarian cancer classification. This indicates in both medical and research areas that can rely on its efficacy to correctly identify ovarian cancer, enhancing the diagnosis of ovarian and facilitating ovarian cancer detection for improved treatment. See table 6.

Table 6: Proposed model K-fold results

Model	K-Fold	K-Fold Accuracy	Standard Deviation
VGG16 and XGBoost	10	98.37%	0.0025

State of the Art Comparison

Table 7 offers a comparative performance assessment of our proposed model alongside current methodologies to deliver a thorough evaluation of its effectiveness. Most current studies mainly depend on machine learning techniques, reaching a maximum accuracy of 98% for detecting ovarian cancer. In comparison, the suggested deep learning model, utilizing a groundbreaking approach greatly exceeds these benchmarks by achieving an outstanding performance result of 99.7%. it is essential to highlight that all previous research categorized a smaller number of ovarian cancer subtypes, while our model successfully classified a greater variety of ovarian cancer subtypes.

Table 7: Proposed model compared with state of the art

Reference	Technique	Model	Accuracy	Classes
Guangxing Wang, 2023	Deep Learning	CNN (fine-tuned based on pretrained ResNet50 framework)	99.7%	3
Anvita Reddy Inture et al, 2023	Machine Learning	Random Forest	98%	2

This analysis clearly shows that the suggested methodology is greatly more effective than the current leading studies. The suggested model surpasses conventional detection methods, greatly enhancing the accuracy and reliability of benchmarks for identifying subtypes of ovarian cancer in comparison to other approaches discussed in the literature. These results demonstrate the strong potential of this hybrid model for identifying ovarian cancer subtypes. By

setting up a higher performance benchmark, the proposed approach offers a meaningful step forward for automated cancer diagnostics. The precision and accuracy of diagnosing ovarian cancer will increase through this approach. The innovative approach creates a reliable tool for oncologists to automate diagnostic processes and lead to improved classification and diagnostic approaches for ovarian cancer.

CONCLUSION

Remarkable classification outcomes of the ovarian cancer system highlight the strength of merging XGBoost as a classification method with VGG16 for extracting features through transfer learning. Analysis of various ovarian cancer types achieved outstanding results through the model which show high performance accuracy across multiple K-fold Cross
validation runs. Through macro-average scoring definitions the model becomes prepared to make predictions across numerous ovarian cancer types. The macro-averaged scores explicitly verify the models ability to generalize its performance. The K-fold cross validation method produced breakthrough performance by reaching an accuracy rate of 98.37% along with a standard deviation of 0.0025.

6 FUTUREDIRECTION

Our future research will explore additional approaches to enhance data augmentation for expanding the sample collection with new techniques to balance datasets alongside the incorporation of extra types of ovarian cancer and performance boost methods. The approach leads to the creation of a more powerful and efficient system for ovarian cancer classification.

REFERENCES

M. A. Bookman et al., "Evaluation of new platinum-based treatment regimens in advanced-stage ovarian cancer," Journal of Clinical Oncology, vol. 27, pp. 1419–1425, 2009.

J. Breen, K. Allen, K. Zucker, P. Adusumilli, A. Scarsbrook, G. Hall, N. M. Orsi, and N. Ravikumar, "Artificial intelligence in ovarian cancer histopathology: a systematic review," NPJ Precision Oncology, 2023.

J. Prat and F. C. on Gynecologic Oncology, "Staging classification for cancer of the ovary, fallopian tube, and peritoneum," International Journal of Gynecology and Obstetrics, vol. 124, no. 1, pp. 1–5, 2014. Epub 2013 Oct 22. PMID: 24219974.

T. Liu, R. Su, C. Sun, and X. Li, "Predicting prognosis of epithelial ovarian cancer with whole slide histopathological images," 2022.

H. Zeng, L. Chen, M. Zhang, Y. Luo, and X. Ma, "Integration of histopathological images and multi-dimensional omics analyses predicts molecular features and prognosis in high-grade serous ovarian cancer," 2021.

F. Nourin and M. S. A. Hashmi, "Osteosarcoma bone tumor detection using deep ensemble transfer learning architecture," Spectrum of Engineering Sciences, pp. 2040–2056, 2025.

B. Ziyambe et al., "A deep learning framework for the prediction and diagnosis of ovarian cancer in pre- and post-menopausal women," Diagnostics, 2023.

M. Akazawa and K. Hashimoto, "Artificial intelligence in ovarian cancer diagnosis," Anticancer Research, pp. 4795–4800, 2020.

G. Wang, H. Zhan, T. Luo, B. Kang, X. Li, G. Xi, Z. Liu, and S. Zhuo, "Automated ovarian cancer identification using end-to-end learning and second harmonic generation imaging," IEEE Journal of Selected Topics in Quantum Electronics, p. 7200609, 2023.

A. M. S., A. I., A. Kodipalli, and R. J. Martis, "Ovarian cancer detection and classification using machine learning," in ICEECCOT, p. 280, 2021.

K. R. Kasture, D. D. Shah, and P. N. Matte, "A new deep learning method for automatic ovarian cancer prediction and subtype classification," Turkish Journal of Computer and Mathematics Education, pp. 1233–1242, 2021.

M. Wu, C. Yan, H. Liu, and Q. Liu, "Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks," Bioscience Reports,
pp. 1–14, 2018.
J. Fan, J. Liu, Q. Chen, W. Wang, and Y. Wu, "Accurate ovarian cyst classification with a lightweight deep learning model for ultrasound images," IEEE Transactions on Medical Imaging, pp. 110681–110682, 2023.
K. R. Kasture, B. B. Sayankar, and P. N. Matte, "Multi-class classification of ovarian cancer from histopathological images using deep learning - vgg-16," in GCAT, pp. 1–10, 2021.
A. Ahmed, Z. Xiaoyang, M. H. Tunio, M. H. Butt, and S. A. Shah, "Occnet: Improving imbalanced multi-centred ovarian cancer subtype classification in whole slide images," in ICCWAMTIP, pp. 1–10, 2023.
A. Krizhevsky, I. Sutskever, and G. E. Hinton, "A deep learning framework for cancer diagnosis using medical imaging," IEEE Transactions on Medical Imaging, 2012.
K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint, 2014.
F. Chollet, Deep Learning with Python. Manning Publications, 2017.
A. G´eron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly Media, 2019.
M. S. A. Hashmi, A. M. Qadri, A. Raza, S. Ullah, A. Smerat, C. Kim, M. Syafrudin, and N.
L. Fitriyani, "Vggbm-net: A novel pixel-based transfer features engineering for automated coffee bean diseases classification," IEEE Access, 2025.
T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in KDD, pp. 785–794, 2016.
K. Haider, M. F. Asghar, M. S. A. Hashmi, M. Saeed, R. Fatima, and S. N. Khosa, "Xcepforest: A hybrid deep learning and machine learning approach for endangered aquatic species classification," Spectrum of Engineering Sciences, pp. 476–485, 2025.
T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning. 2 ed., 2009.
A. G. Howard et al., "Searching for mobilenetv3," in ICCV, pp. 1314–1324, 2019.
Y. LeCun, F. J. Huang, and L. Bottou, "Learning methods for generic object recognition with invariance to pose and lighting," in CVPR, 2004.
C. Nebauer, "Evaluation of convolutional neural networks for visual recognition," IEEE Transactions on Neural Networks, vol. 9, no. 4, pp. 685–696, 1998