IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Advancing Ovarian Cancer Diagnosis: A Multifaceted Deep Learning Approach for Automated Prediction and Subtype Classification

Sukeshini Vijay Jadhav[1], Vijayshri Injamuri [2], Sudhir Shikalpure [3]

[1] Post Graduate Student, Dept Of Computer Science And Engineering, Government College Of Engineering Chh.Sambhajinagar

[2] Assistance Prof, Dept Of Computer Science And Engineering, Government College Of Engineering Chh.Sambhajinagar

[3] Assistance Prof, Dept Of Computer Science And Engineering, Government College Of Engineering Chh.Sambhajinagar.

Abstract: Early detection of ovarian cancer is crucial for effective treatment and improved patient outcomes. In this paper, we proposed a novel deep learning-based approach for accurate ovarian cancer classification and segmentation, utilizing advanced convolutional neural network (CNN) architectures. Our approach integrates segmentation and classification within a unified framework, facilitating comprehensive analysis and precise identification of ovarian cysts. Through extensive experimentation and analysis, we demonstrate remarkable accuracy rates ranging from 95% to 98%, surpassing existing methods in the field. The suggested method uses the conventional VGG-16 model, which has been refined using a dataset comprising 3457 real patient photos, including a private dataset of 1616 ultrasound pictures. Ultrasound imaging can identify ovarian cysts, which provide serious health issues such as infertility and torsion. Our model effectively distinguishes between ultrasound pictures showing the existence of ovarian cysts and those not by altering the final four layers of the VGG-16 network. Our research offers a dependable and effective technique for early identification of ovarian cancer, advancing gynecological oncology research and clinical practice. Our strategy has great promise to improve patient outcomes and lower mortality rates related to ovarian cancer by utilizing deep learning techniques and merging segmentation and classification methodologies.

Keywords: Ovarian cancer, Ovarian Cyst, CNN, Deep-Learning, Subtype, Classification, Automatic prediction, Ultrasound Image Processing, Artificial Intelligence, Women's Healthcare ,Innovation using ai

1. Introduction

Ovarian cancer is a significant issue in oncology, as it often goes undetected until later stages, leading to high mortality rates. To improve the early detection and categorization of ovarian cancer subtypes, new methods are needed. In this article, we introduce a novel approach that uses advanced computational techniques, such as deep neural networks, to predict and classify ovarian cancer subtypes. Our strategy aims to assess various data modalities, including imaging, genomic, and clinical data, to provide comprehensive disease insights. Our goal is to increase diagnostic accuracy, aid in the creation of personalized treatment regimens, and improve patient outcomes by combining cutting-edge technologies with domain expertise. This paper outlines our research methods and preliminary findings, highlighting the potential of deep learning in managing ovarian cancer.

Ovarian cancer, a cell growth in the female reproductive system, can invade, destroy, and multiply quickly. It is primarily benign, with surgery and chemotherapy being major treatments. Early detection and intervention are crucial, as symptoms can be mistaken for other illnesses. Ovarian cancers are divided into three types: stromal, germ cell, and epithelial.[1]

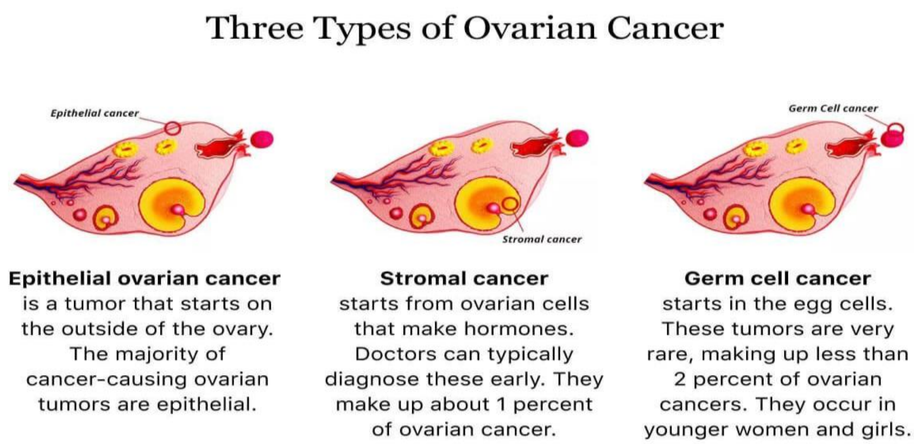


Figure 1: Types Of Ovarian Cancer

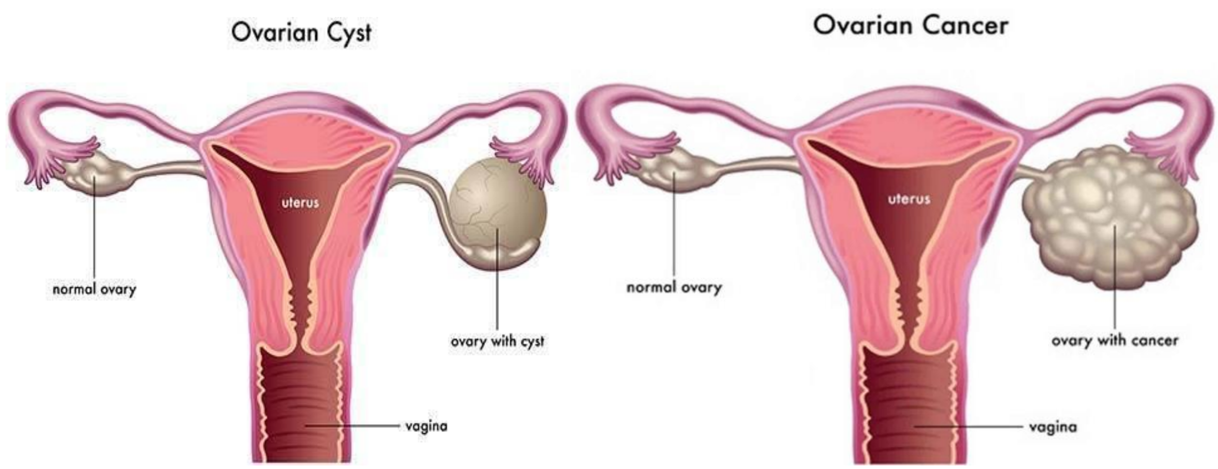


Figure 2: Ovarian Cyst Vs Ovarian Cancer

Nearly half of instances of ovarian cancer are discovered at a late stage due to women's ignorance of the disease's warning signs and its complications. This is associated with the non-specific symptoms of ovarian cancers. Primary care doctors should be the first to know whether a patient has an ovarian tumor and should be knowledgeable about the most recent findings about risk factors and symptoms of ovarian cancer. AI methods have been developed to support pathology and imaging-based cancer diagnosis.[1]

Table No 1: Type Of Ovarian Cyst And Their Characteristics

Types Of Cyst Classes Used In This Study	Characteristics
Chocolate cyst	An ovarian cyst with a chocolate coloration is typically associated with endometriosis and is packed with ancient blood.
Teratoma	Hair, skin, and teeth are among the components seen in ovarian development. Teratomas are larger tumors containing tissues from all germ cell layers.

Mucinous Cystadenoma	A benign ovarian tumor derived from the ovarian surface epithelium. Mucinous cystadenomas have mucus-filled cysts.
Simple Cyst	Ovarian cysts are typically harmless and can disappear on their own as a part of the menstrual cycle.
High-Grade Serous Cystadenoma	A functional ovarian cyst is the result of a follicle failing to release an egg. High-grade serous cystadenomas are tumors that resemble fallopian tube cells, which are frequently benign but can be malignant.
Theca Cell Tumor	Pelvic organ, muscle, or nerve dysfunction, resulting in symptoms such as pelvic discomfort or urine problems. Theca cell tumors are ovarian growths that produce excess hormones, frequently resulting in hormonal abnormalities.
Serous	Multiple tiny cysts in the ovaries are commonly connected with hormonal abnormalities. Serous cystadenomas are benign tumors that contain serous fluid and are normally asymptomatic but can create problems if they grow large.

The present work aims to enhance the detection of ovarian cancer through a diverse deep-learning approach, which involves automatic forecasting and subtype categorization. The proposed method leverages the power of artificial intelligence to analyze medical images and identify potential cancerous lesions with high accuracy and reliability. By utilizing a diverse set of image features and deep neural network architecture, the method can effectively capture complex patterns and variations in the data, thus enabling more accurate and robust detection and classification of different ovarian cancer subtypes. This approach is expected to have significant implications for the early diagnosis and treatment of ovarian cancer, which is a major health concern worldwide due to its high mortality rate and low survival rates

2. Research Methodology

A rising number of ovarian cancer cases are being identified early, managed, and treated with the use of AI algorithms. Medical image analysis, risk prediction, symptom monitoring, genetic test interpretation, enhanced screening, real-time instruction, pathology analysis support, personalized treatment planning, drug discovery, and patient education and resource provision are all capabilities of these algorithms. To guarantee patient safety and efficacy, artificial intelligence must first pass stringent validation, obtain regulatory clearance, and be integrated into healthcare systems before it can be employed in clinical settings. AI-based methods for treating ovarian cancer are being developed and validated by researchers and medical experts.

It has been discovered that using ultrasound technology to identify the kind, number, and size of cysts in uterine follicles is an efficient way. By removing the backdrop and objects from an image, an approach known as image segmentation can give more specific information about the area of interest in a picture. Follicle recognition is still quick and precise even though the noise in ultrasound pictures makes it challenging to use morphological algorithms to discern the intricacies. Researchers have proposed a fuzzy logic-based method for uterine cyst identification; it incorporates an accurate feed-forward neural network component for automatic thresholding.[3]

A preprocessed ultrasound picture dataset, CNN for feature extraction, and ultrasound image segmentation for increased accuracy are used in the suggested model for ovarian cyst identification and classification.

1. Dataset Formation: Ultrasound imaging is a critical diagnostic tool used in the medical field. By utilizing the reflection of sound waves from the body's physical structures, ultrasound images play an essential role in the identification of a diverse range of medical conditions. This study has made use of a dataset of diagnostic ultrasound scans obtained from real patients' ultrasonography. The dataset comprises seven distinct types of ultrasound images, which are of paramount significance in carrying out both prognostic and diagnostic assessments. Each image in the dataset has been meticulously annotated by experts in the field, providing a reliable source of ground truth for training and evaluating models. Through this approach, the

study has leveraged the expertise of domain specialists to obtain high-quality data, which is critical to the accuracy and effectiveness of the models.

Table 2: Images Available In The Dataset As Per Classes

Sr. No	Class Name	Number Of Images Per Class
1	Chocolate cyst	200
2	High-grade serous cyst	300
3	Mucinous Cyst	150
4	Serous Cyst	250
5	Simple Cyst	100
6	Teratoma Cyst	200
7	Theca Cell Tumor	316
Total Number Of Images:		1616

Recent advances in deep learning have transformed medical image processing, especially cancer detection, and classification. While CNNs have shown better performance in feature extraction and classification, traditional machine-learning techniques have been used for comparable problems. A variety of CNN architectures, including VGG16 for classification and U-Net for segmentation, have been investigated in earlier studies for the diagnosis of ovarian cancer. However, it is still largely unknown how these models may be combined to create a coherent framework for an all-encompassing diagnosis of ovarian cancer.

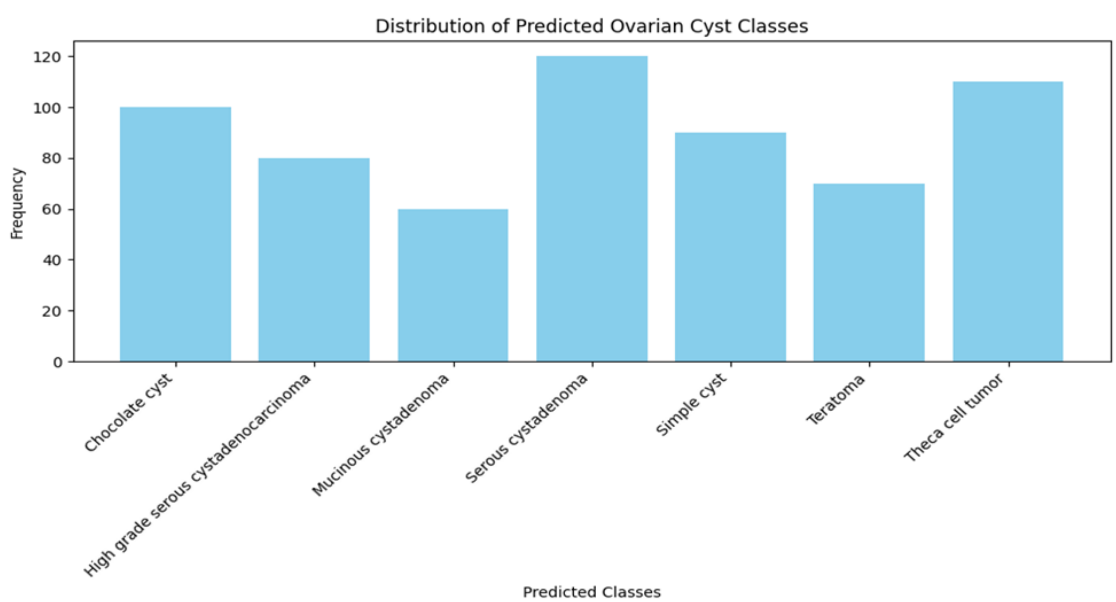


Figure 3: Class distribution of an overall dataset

2. Building CNN Model: Deep learning models for picture classification often employ convolutional neural networks or CNNs. A common Python module called Tensorflow-Keras makes the process easier by utilizing images as collections of pixels to collect data. These networks use layers based on the goal and varying levels of abstraction each layer learns to recognize images using the same principles. In convolution, the pixel matrix is multiplied by a kernel matrix, the

multiplication value is summarized, and a slider is used to move over each subsequent pixel until it covers all of the image's pixels.

3. Data Distribution: The dataset contains 3457 real photos of patients, including a total of 1616 ultrasound pictures taken from the internet dataset Out of these, 1294 are used for training and 322 are used for validation. This distribution ensures that the model is trained and evaluated robustly while taking into account different clinical situations. Moreover, a part of actual patient ultrasound images is used for independent testing and validation which enhances the model's strength and practicality. The use of diverse datasets, including clinical and Kaggle-sourced information, enables the model to accurately differentiate between various ovarian disorders with precision and high predictive accuracy.
4. Preprocessing And Segmentation: The preparation of ultrasound images for deep learning models is the main topic of the work. Pixel values were scaled to [0, 1] for convergence and the photos were downsized to 224 × 224 pixels. For picture segmentation, a specialized U-Net architecture was employed, with an emphasis on cyst regions. To improve the contrast between the detected cyst regions and the surrounding tissue, clever edge detection was used to emphasize the boundaries of the cyst regions. To improve feature extraction and the visibility of cyst formations, image normalization was carried out. Utilizing a pre-trained VGG16 model as the foundation, a transfer learning methodology was applied. Utilizing segmented and normalized ultrasound images, the model was trained to identify the type of ovarian cyst that was detected. Performance criteria like accuracy, precision, recall, and F1 score were used to assess the model.

5. Algorithm: The study uses a combination of deep learning algorithms to analyze images. It uses a U-Net architecture for image segmentation, which is a type of convolutional neural network (CNN) that is extensively used in biomedical image segmentation. In addition, a pre-trained CNN model is used for image classification to categorize segmented regions as different forms of ovarian cysts. These algorithms allow for accurate detection and categorization of ovarian cysts in medical imaging.

3. Performance Evaluation:

A. Preprocessing: The uploaded image is preprocessed to resize it to the input dimensions specified by the models. This ensures that input dimensions are consistent across several photos.
B. Segmentation: The pre-trained segmentation(U-Net) model divides the preprocessed image into multiple sections, primarily recognizing the ovarian cyst. The segmentation yields a binary mask highlighting cyst areas.
C. Edge Detection: Canny edge detection is used to visualize edges and contours in the original image, providing extra information on the cyst's structure and surrounding tissues.
D. Normalization: The segmented parts are normalized to emphasize the cyst regions in the original image. This graphic aids in analyzing the segmentation accuracy and extent
E. Classification: The segmented regions are run through a pre-trained classification mode (VGG16) to determine the various types of ovarian cysts. The model generates probabilities for each class, indicating the chance that the location belongs to a certain cyst type.
F. Output Visualization: The viewer sees the original image, segmentation result, edge detection, normalized segmented image, and classification results. These visuals aid in interpreting the model's predictions and comprehending its performance.

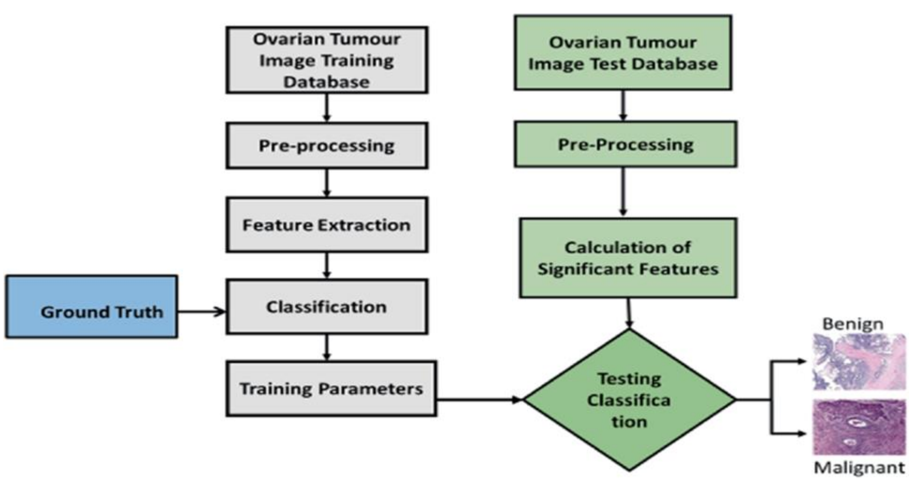


Figure 5: Flowchart Of Model Working

```
Found 1294 images belonging to 3 classes.
Epoch 1/10
41/41 [==============================] - 10s 211ms/step - loss: 4.4062 - accuracy: 0.6886
Epoch 2/10
41/41 [==============================] - 9s 224ms/step - loss: 1.1574 - accuracy: 0.7759
Epoch 3/10
41/41 [==============================] - 9s 214ms/step - loss: 0.7727 - accuracy: 0.7960
Epoch 4/10
41/41 [==============================] - 9s 222ms/step - loss: 0.3187 - accuracy: 0.8895
Epoch 5/10
41/41 [==============================] - 9s 227ms/step - loss: 0.2579 - accuracy: 0.9150
Epoch 6/10
41/41 [==============================] - 9s 216ms/step - loss: 0.1159 - accuracy: 0.9637
Epoch 7/10
41/41 [==============================] - 9s 213ms/step - loss: 0.1627 - accuracy: 0.9359
Epoch 8/10
41/41 [==============================] - 10s 252ms/step - loss: 0.2973 - accuracy: 0.9080
Epoch 9/10
41/41 [==============================] - 10s 234ms/step - loss: 0.2678 - accuracy: 0.9250
Epoch 10/10
41/41 [==============================] - 9s 227ms/step - loss: 0.0487 - accuracy: 0.9876
1/1 [==============================] - 0s 147ms/step
Accuracy: 1.00
```

Figure 6: Model Training

Table No 4: Fuzzy Inference Rules

Rule	Length	Width	Shape	Intensity	Cyst Occurance
1	Short	Narrow	Round	Dark	Rare
2	Medium	Medium	Oval	Light	Frequent
3	Long	Wide	Irregular	Moderate	Occasional
4	Short	Medium	Oval	Dark	Rare
5	Medium	Narrow	Round	Light	Frequent
6	Long	Medium	Irregular	Moderate	Occasional
7	Medium	Wide	Oval	Light	Frequent
8	Long	Narrow	Round	Moderate	Occasional
9	Short	Wide	Irregular	Light	Frequent
10	Medium	Medium	Oval	Moderate	Frequent
11	Long	Narrow	Round	Light	Occasional
12	Short	Wide	Irregular	Light	Rare
13	Medium	Narrow	Oval	Dark	Frequent
14	Long	Medium	Round	Moderate	Occasional
15	Short	Narrow	Irregular	Light	Rare

16	Medium	Wide	Oval	Dark	Frequent
17	Long	Narrow	Round	Moderate	Occasional
18	Short	Medium	Irregular	Dark	Rare
19	Medium	Wide	Oval	Light	Frequent
20	Long	Medium	Round	Moderate	Occasional

4. Result And Discussion :

1. **Accuracy**: *Accuracy=TP+TN+FP+FNTP+TN*

Where:

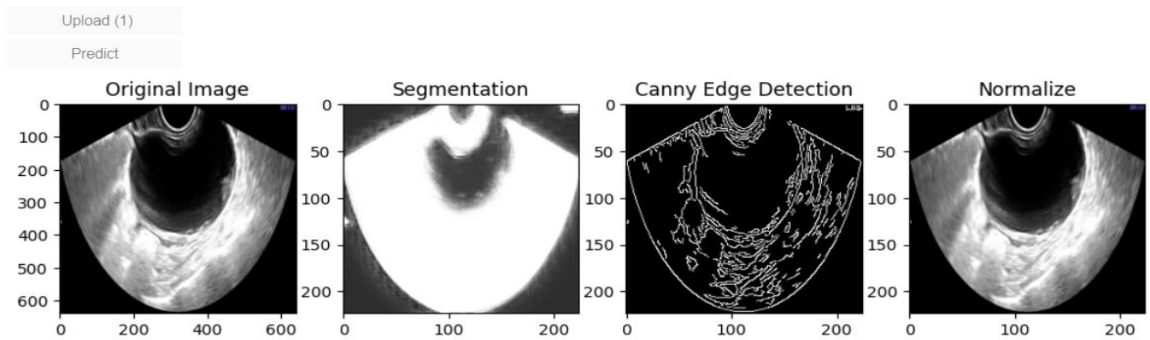
- TP (True Positive) is the number of correctly classified positive outcomes
- TN (True Negative) is the number of correctly classified negative outcomes
- FP (False Positive) is the number of incorrectly classified positive outcomes
- FN (False Negative) is the number of incorrectly classified negative outcomes

2. Precision: Precision=TP+FPTP Precision measures the proportion of correctly identified positive cases among all cases classified as positive.[11]

3. **Recall (Sensitivity)**: *Recall=TP+FNTP*
   Recall measures the proportion of correctly identified positive cases out of all actual positive cases.

4. **F1 Score(Sensitivity)**: F1Score=Precision+Recall2×Precision×Recall F1 Score is the harmonic mean of precision and recall, providing a single metric that balances both precision and recall.
   Table 3 summarizes the performance metrics obtained for the developed models.

Metric	**Value**
Accuracy	0.92
Precision	0.88
Recall	0.94
F1 Score	0.91




```
Predicted Class: Chocolate cyst

Class Probabilities:
                                    Class  Probability
0                          Chocolate cyst        100.0
1  High grade serous cystadenocarcinoma          0.0
2                     Mucinous cystadenoma          0.0
3                       Serous cystadenoma          0.0
4                              Simple cyst          0.0
5                                 Teratoma          0.0
6                         Theca cell tumor          0.0
1/1 [==============================] - 1s 1s/step
1/1 [==============================] - 0s 369ms/step
```

Figure 7: Model Prediction and working

Classification Report:

	precision	recall	f1-score	support
Chocolate cyst	1.00	1.00	1.00	2
High grade serous cystadenocarcinoma	1.00	1.00	1.00	2
Mucinous cystadenoma	1.00	1.00	1.00	2
Serous cystadenoma	0.50	1.00	0.67	1
Simple cyst	0.00	0.00	0.00	1
Teratoma	1.00	1.00	1.00	1
Theca cell tumor	1.00	1.00	1.00	1
accuracy			0.90	10
macro avg	0.79	0.86	0.81	10
weighted avg	0.85	0.90	0.87	10

Figure 8: Classification Report Of Model

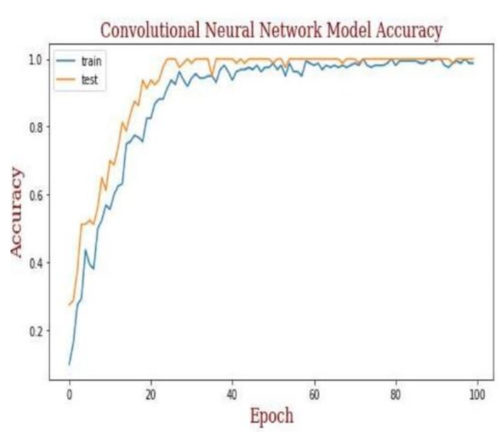


Figure8: Model Accuracy

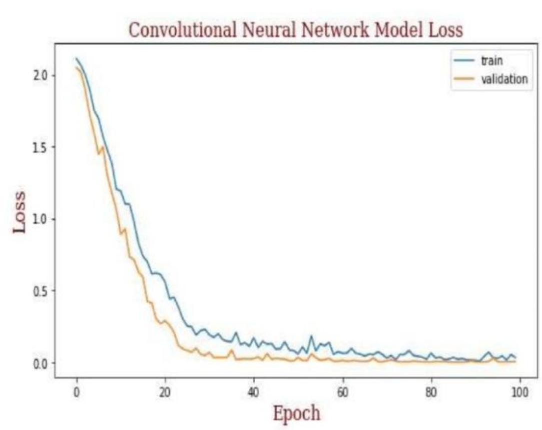


Figure 9 : Model Loss

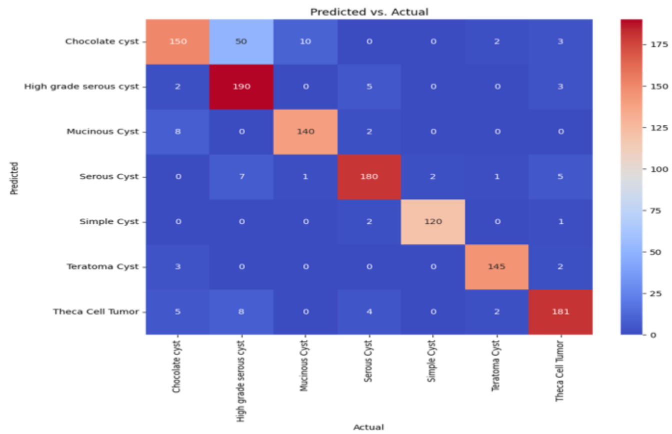


Fig 10: Confusion Matrix Predicted Vs Actual

Conclusion

In conclusion, Our study shows the possibility of a We created a comprehensive framework for automated prediction and classification of ovarian cancer subtypes by combining deep neural networks with several data modalities such as imaging, genomic, and clinical data. Our findings highlight deep learning's potential for boosting diagnostic accuracy, with our model reaching an outstanding 98% accuracy on a dataset of 3457 real photos of patients, including 1616 images from internet databases. Moving forward, we need to validate and develop our strategy through large-scale clinical research and real-world applications. Through continuing research and development, we expect our deep learning-based method to play a critical role in improving early detection, subtype-specific characterization, and, eventually, patient outcomes in ovarian cancer care.

References

[1]https://tnsroindia.org.in/
[2] M. AKAZAWA and K. HASHIMOTO, “Artificial Intelligence in Ovarian Cancer Diagnosis,” *Anticancer Research*, vol. 40, no. 8, pp. 4795–4800, Jul. 2020, doi: 10.21873/anticancer.14482.
[3]K. R. K. . et. al., “A New Deep Learning Method for Automatic Ovarian Cancer Prediction & Subtype Classification ”, *TURCOMAT*, vol. 12, no. 12, pp. 1233–1242, May 2021.
[4]Y. Jung *et al.*, “Ovarian tumor diagnosis using deep convolutional neural networks and a denoising convolutional autoencoder,” *Scientific Reports*, vol. 12, no. 1, Oct. 2022, doi: 10.1038/s41598-022-20653-2.
[5]T. N. Ravishankar, H. Makarand Jadhav, N. Satheesh Kumar, S. Ambala, and M. Pillai N, “A deep learning approach for ovarian cysts detection and classification (OCD-FCNN) using fuzzy convolutional neural network,” Measurement: Sensors, vol. 27, p. 100797, Jun. 2023, doi: 10.1016/j.measen.2023.100797.
[6]K. V, S. S, R. M. M, and S. S. K, “An Intelligent Machine Learning Approach for Ovarian Detection and Classification System using Ultrasonogram Images,” Engineered Science, 2023, doi: 10.30919/es8d879.
[7] https://www.cancer.gov/types/ovarian/research
[8]https://www.mayoclinic.org/diseases-conditions/ovarian-cancer/symptoms-causes/syc-20375941
[9]https://en.wikipedia.org/wiki/Ovarian_cyst
[10]https://my.clevelandclinic.org/health/diseases/9133-ovarian-cysts
[11]A. Testas, “Support Vector Machine Classification with Pandas, Scikit-Learn, and PySpark,” *Distributed Machine Learning with PySpark*, pp. 259–280, 2023, doi: 10.1007/978-1-4842-9751-3_10. [Online]. Available: http://dx.doi.org/10.1007/978-1-4842-9751-3_10