A Convolution-Attention Framework with Explainable AI for Ovarian Cancer Diagnosis using Histopathology Images

Hrishita Panjetha[1], Jitendra Singh Jadon[2]
{hrishita.panjetha@gmail.com[1], jsjadon@amity.edu[2]}

Amity University, Noida[1], Amity University, Noida[2]

Abstract. Ovarian cancer has been one of the deadliest gynecologic cancers whose survival of the patient directly depends on the initial and accurate diagnosis measure. This study presents a Convolution-Attention architecture where Explainable Artificial Intelligence (XAI) technology is applied to the automation, multi-class categorization of ovarian cancer with the help of histopathological images. The main model is also based on the CoAtNet architecture that combines convolutional operations to extract local features and self-attention networks to extract the global dependencies of images. The system was trained and tested using a collection of 5 histopathological categories namely Clear Cell, Endometrioid, Mucinous, Serous, and Non-cancerous ovarian tissues. Emprical results reveal test reliability of 98.26 percent and a macro-average of F1-score of 0.9823 and thus outperforms the known CNN-based low-level models. Essentially, interpretability was also obtained using Grad-CAM visualisations which explain diagnostically meaningful tissue areas and ensure consistency between model outputs and typical pathological structures. The intersection of learning deep features and possibility of transparent visual explanations confirms the validity of the presented framework as reliable and understandable, which provides a reliable decision support system to clinical pathologists. This paper highlights the opportunities offered by having hybrid convolution-attention networks that have been augmented with explainable AI to improve the accuracy of the overall ovarian cancer diagnostics.

Keywords: Deep learning, CoAtNet, Explainable Artificial Intelligence, Grad-CAM, Medical image classification, Convolution-Attention networks

1 Introduction

As violating most of the known rules of cancer invasion, ovarian cancer is still among the deadliest gynecological cancers in the world due to its nature of progressing without any symptoms and detection at an advanced stage of the disease. Timely diagnosis stands as a key factor to enhance the survival rate but credible diagnosis procedures in clinical practice are still hard to come by.

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2362623

Histopathologic study of tissue slides is the existing gold standard of diagnosis of ovarian cancer to date[9], as it provides microscopic evidence on the morphology of tumors, and helps to characterize the subtypes. Nevertheless, manual analysis is not only time-consuming and subjective, but also can be affected by inter-observer variability, which highlights the necessity of automated and objective support systems of diagnostic assistance.

The recent developments in artificial intelligence (AI)[7] and deep learning have demonstrated outstanding prospects in the field of computational pathology. Convolutional neural networks (CNNs) have been shown to be capable of learning discriminative image features to detect cancer and subtype within a single stage of learning, and in the same stage, learning to discriminate between healthy and tumorous tissues with high accuracy and success, as well as, directly in the test data grouping and localizing tumors as subgroups relative to normal tissues[20]. However, the CNNs often have been faced with restrictions in modeling global spatial dependencies that are imperative in the understanding of the complex tissue structures in histopathology. To bridge this gap, hybrid architectures that combine convolutional and attention mechanisms have emerged as a powerful approach for medical image analysis, offering both efficiency and contextual understanding.

The current study presents a Convolution-Attention architecture based on the CoAtNet [2] architecture on the multiple-classification of subtypes of ovarian cancer through histopathology. The architecture uses convolutional layers to extract local features but uses attention to extract long-range dependencies and therefore, enable a holistic representation of the tissue morphology. The research uses a set of data that includes five classes, such as Clear Cell, Endometrioid, Mucinous, Serous, and Non-cancerous ovary tissues. The training pipeline is designed to achieve robustness and reduce the variances of classes with the help of large-scale data augmentation plans such as rotational transformations, horizontal and vertical flipping, and additive noise.

One of the main novelties offered in this paper is the introduction of Explainable Artificial Intelligence (XAI)[17] methods to make predictions of the model interpretable. Grad-CAM visualisation is employed to map out the parts of the image that have a strong effect on the classification performance to provide essential information regarding the circuitry in the network that makes decisions and build upon clinician trust.

Although the research of hybrid convolution-attention systems and explainable artificial intelligence has been explored individually, the presented study presents a purposely designed system that optimizes on ovarian cancer histopathology in particular. The proposed architecture (i) uses CoAtNet variant derived after empirical investigations of the complexity of features in ovarian tissue, (ii) dynamic baseline hierarchical convolution attention block with a domain specific augmentation strategy to enhance the magnitude of the glandular and stromal variability, and (iii) integrates class -consistent Grad-CAM reasoning mechanism integrating attention with previously validated morphological features. Such is an intentional design that contrasts the model with the previous deep-learning systems, which have either no global-local feature fusion or a lack of clinical alignment in their qualitative interpretability.

Overall, it is expected that the suggested Convolution-Attention model can achieve high diagnostic accuracy and remain interpretable, thus bridging the gap between performance and interpretability of a study. This approach promotes AI-supported ovarian cancer diagnostics, and the possibility of hybrid deep-learning paradigms is highlighted to support the work of the pathologist

with reliable, reproducible, and timely disease detection.

2 Literature Survey

Most of the studies have investigated the use of machine-learning and deep-learning approaches to improve the diagnostic and prognostic evaluation of ovarian cancer in recent studies. Lu et al. (2020)[5] used a machine learning framework that is based on decision-trees and merged Minimum Redundancy Maximum Relevance (MRMR) feature selection to differentiate between benign and malignant tumours in the human ovaries, achieving an accuracy of 92.1, and explaining the results by critical biomarkers like HE4 and CEA, among other biomarkers[5]. The study by Zhang et al. (2023)[6] on a tumor-infiltrating immune-cell-related pseudogene signature (TIICPS) was built by combining ten different machine-learning algorithms on three genomic datasets, providing the concordance index (C-index) of 0.749 and outstanding potential predictors of immunotherapy outcomes in the case of ovarian cancer patients. Kasture et al. (2022)[4] introduced KK -Net, which is a higher-quality Deep Convolutional Neuron Network (DCNN) that is trained on boosted histopathology images of the TCGA-OV and achieved high classification rates of up to 91% compared to existing CNN models like AlexNet and VGGNet.

A more comprehensive exploration of the classification of ovarian cancer was done by Radhakrishnan et al. (2024)[1] who provided a comparative study on various deep-learning architectures that are the state of the art such as MobileNetV2, VGG19, ResNet18, Xception, EfficientNet, and InceptionV3. Their experimental results identified InceptionV3 as the top-performing model, achieving an accuracy of approximately 97.96% on their multiclass ovarian cancer dataset. It is a factor that set this study apart since it focused on model interpretability and used various explainability tools to visualize the discriminative features that informed each prediction such as: Integrated Gradients, Saliency Maps, Grad-CAM, and DeepLIFT. These interpretive layers did not only help to increase transparency, but provided a useful clinical validation tool. However, the paper was limited by dataset-specific appraisals and had not been analyzed externally, which is a requirement to determine the suitability of the of diagnostic AI systems.

A systematic review of a machine-learning based algorithm to predict and diagnose ovarian-cancer was undertaken by Kwatra et al. (2023)[10], which identified the Support Vector Machines (SVM), Random Forests, Logistic Regression models with a maximum accuracy of 97.86% on benchmark data sets. A comparative analysis is shown in Table 1.

3 Dataset Description

As a dataset, the current study uses Ovarian Cancer and Subtypes Histopathology dataset, initially used by Meera Radhakrishnan et al. and consists of the image patches that are hematoxylin and eosin (H8E)-stained, and correspond to five different categories: Clear Cell, Endometrioid, Mucinous, Serous, and Non-cancerous of the ovarian tissues. The dataset is comprised of 499 high-resolution histopathological images, which were cut out of whole-slide scans and then standardized to a resolution of 224 224 pixels [11]. Each image was stain-normalized and augmented through random rotations, flips, and brightness adjustments to enhance generalization. This dataset provides

Table 1: Summary of Related Studies in Ovarian Cancer Diagnosis

Author and Year	Model Used	Accuracy (%)
Lu et al. (2020)	Decision Tree with MRMR feature selection for distinguishing benign and malignant ovarian tumors.	92.1
Zhang et al. (2023)	TIICPS hybrid model integrating ten machine learning algorithms for immune-related prognostic prediction.	74.9
Kasture et al. (2022)	Enhanced Deep CNN (KK-Net) trained on augmented TCGA-OV histopathology images, outperforming traditional CNNs.	91.0
Kwatra et al. (2023)	Comparative study of ML algorithms (SVM, Random Forest, Logistic Regression) for ovarian cancer prediction.	97.8
Radhakrishnan et al. (2024)	InceptionV3 (Best among multiple CNN backbones).	97.96

a diverse yet balanced representation of ovarian cancer subtypes, enabling effective evaluation of deep learning models for automated histopathological classification.

4 Data Augmentation

Large scaled data augmentation was applied systematically to the ovarian cancer histopathology dataset to increase generalization of the model and overcome overfitting. The limited coverage of annotated medical images warrants the addition of variation to simulate what is seen in real life, including tissue orientation, lighting differences, and staining variations.

A carefully edited series of geometric and photometric transformations[19] was applied to each image in order to diversify the data set without losing the irreplaceable morphological characteristics of an image. The transformations used included the 15 degree rotations (anti-clockwise and clockwise), horizontals and vertical inversion, zoom of 1.155 and shear of 10° to apply affine distortions. These spatial augmentations were complemented by Gaussian noise of standard deviation 12 and salt-and-pepper noise of corruption ratio 2% in order to mimic imaging artifacts and variability in microscopic sensor.

This was done through methods like Python Imaging Library (PIL) and NumPy, which ensured that there was consistency between two batches of images. The augmented samples were stored together with their original image in order to maintain a reproducible dataset structure. The dataset was therefore increased to more than eight-fold in size, which had the effect of providing the model with a highly enriched exposure to different feature patterns between subtypes of ovarian tissues.

Figure 1 illustrates exemplar augmentations that are presented in various histopathology images. The number supports the principles of geometric and noise transformations developing controlled variability and preserving critical cellular and glandular attributes. The suggested CoAtNet

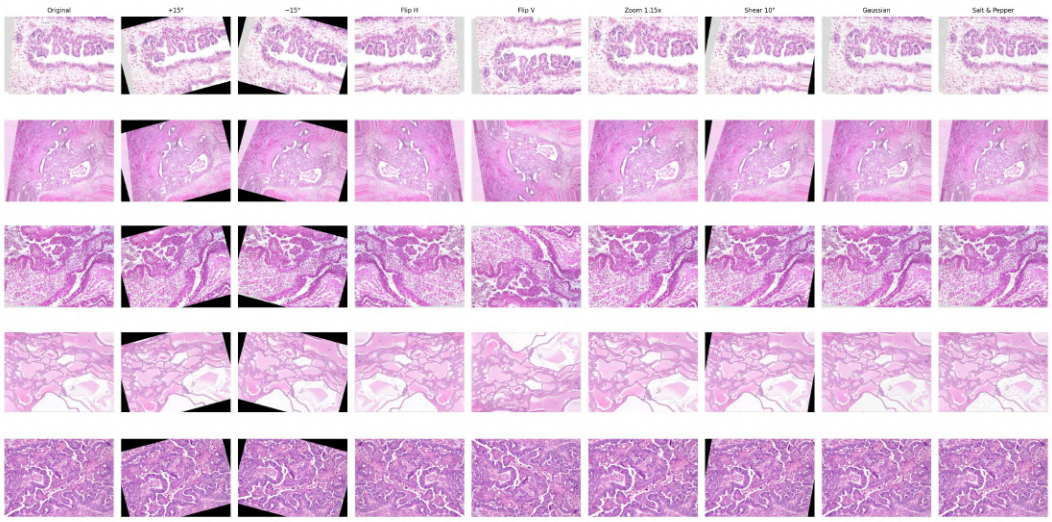


Fig. 1. Sample of Augmented Dataset

model was enhanced by this augmentation pipeline to obtain convincing representations that promoted the robustness as well as classification effectiveness of the proposed pipeline in the five types of ovarian cancer.

Although the dataset was almost balanced between the classes, an imbalance was avoided with (i) class-stratified splitting, (ii) augmentation intensity scaled according to the frequency of classes, augmented and (iii) a class-weighted loss was incorporated during training. These steps were taken to ensure that no subtype had an overbearing impact on gradient updates and thus reduce bias on minority classes like the Serous carcinoma.

5 Methedology

5.1 Model Used

The suggested model is based on CoAtNet (Convolutional and Attention Network)[2] scheme, whereby, the salient features of convolutional neural networks (CNNs) are skillfully batched with Transformer-based attention networks. CoAtNet adopts a hierarchy in an organisation that switches between the convolutional and the self-attention block, thus balancing the local morphological features extraction and global contextual relation inference [3]. Convolutional stages (C1-C2) are to extract small scale cellular information like cell contours and small-scale changes in texture, whereas the attention stages (T1-T2) model large scale spatial to how morphological forms across tissue domains.

This research used CoAtNet0 variant (coatnet 0 rw 224) of the library *timm* which was selected because it is lightweight in computing tasks but has a large representational ability. This model architecture includes stacked depthwise separable convolutions with relative self-attention modules, which are connected by use of feed forward layers and normalisation modules. The head of the final classification was replaced by a fully connected layer with five output neurons associated with the classes that include: Clear Cell, endometrioid, Mucinous, Serous and Non-cancerous. The end-to-end training was done with the Adam optimiser with a learning rate equal to 0.001 using the cross-entropy loss. Early stopping and checkpointing were also added to prevent overfitting. This architecture is a hybrid between the two architectures, which enables it to exploit spatial cues of the low level and semantic relations of high level making it especially useful in complex histopathological image classification problems.

5.2 Dataset Split

The data set was divided into three subsets with 80:10:10 ratio, which were the training, validation, and test respectively. Class stratification was used to ensure that each subset is well represented by all the subtypes. An extensive data augmentation pipeline was enforced to support model generalisation and overfitting prevention; a series of rotations (both clockwise and anti-clockwise, range of -15 to +15), horizontal and vertical flips, and zooming (1.15 times) and shearing (angle of 10 to 10) coupled with manipulative introduction of Gaussian and salt 5p noise. This augmentation method multiplied the dataset more than 8 times and added a realistic variation that could simulate histopathological variation like lighting, staining artefacts, and staining orientation.The intensity of augmentation was determined empirically to reflect realistic histopathological variation without distorting cellular morphology. Higher-intensity transformations introduced artifacts detrimental to classification accuracy, while the selected range preserved structural integrity during training.

5.3 Training and Evaluation Strategy

PyTorch was used to train the model with the augmented dataset with acceleration by a graphical processor to increase the effectiveness of the computations. Mini-batch gradient descent was used and the batch size used was 16 and an adaptive learning-rate schedule was used. In order to obtain robustness, weight regularisation methods and dropout were used, which lowered the overfitting and enhanced the overgeneralization on unseen samples.

After training, the model was tested on the held-out test set. The most important evaluation factors included accuracy, precision, recall and F1-score calculated per-class and also calculated macro-averaged across classes. The test accuracy of the proposed model was 98.26% revealing high levels of discriminative power of the model across all forms of ovarian cancer as shown in high per-class F1-scores (0.97060.9962). These results highlight the ability of the model to learn both morphological and contextual attributes which are essential in proper histopathological distinction.

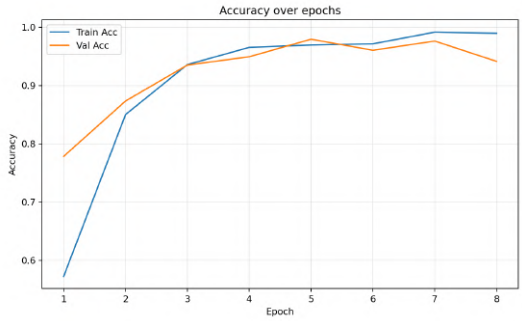


Fig. 2. Training and validation accuracy over epochs.

6 Results

The performance of the proposed Convolution–Attention Framework with Explainable AI was evaluated on the Ovarian Cancer and Subtypes Histopathology dataset, consisting of five categories: Clear Cell, Endometrioid, Mucinous, Serous, and Non-cancerous tissues. The model achieved high classification accuracy and robust generalization across all classes, demonstrating its effectiveness in histopathological image recognition.

6.1 Training Behavior

The convergence in the process of the training was stable in eight epochs. The training and validation loss reduced very fast in the first epochs and leveled off towards the end of the sixth epoch. There is a low variance between the curves, which point to the lack of overfitting. In line with this, trends in the accuracy of Figure 2 reveal steady increment in accuracy of the training process in each epoch with the training accuracy rising to 57.25% in the first epoch, and gradually reaching 98.9% in the last epoch and the validation accuracy contributing to the highest 97.94 %. Learning rate schedule depicts that the initial curve moves exponentially with the exponential scale of 3×10^{-4} to 1.5×10^{-4} which stabilizes optimisation and facilitates a smooth convergence process.

6.2 Quantitative Evaluation

The test accuracy of the model was 98.26% on the held-out test set, macro-average F1-score 0.9823 and weighted F1-score 0.9826 as shown in Table 2. The Endometrioid subtype produced the best F1 -score (0.9962) and the Non-cancerous class maintained a high degree of precision (0.9635) despite a small number of misclassifications.

The model has a discriminative performance as demonstrated in Figure 3, and most of the predictions occur along the diagonal. There were uncommon cases of misclassification and these were mainly under the Mucinous and Non-cancerous category, which could be explained by similarities in the glandular anatomies. These results testify to the fact that the suggested CoAtNet architecture

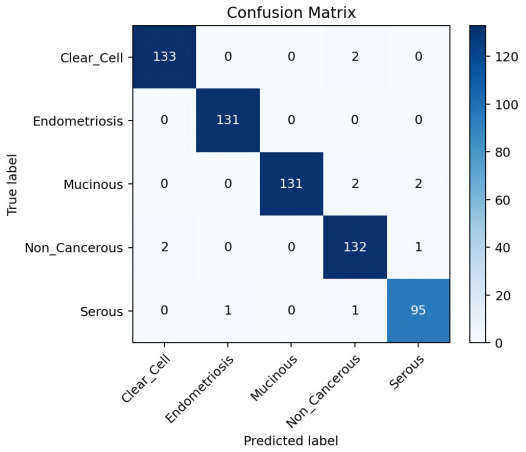


Fig. 3. Confusion matrix for the five ovarian cancer classes.

has good robustness in capturing both textural features at a very fine level as well as higher-order contextual requirements that are most essential in distinguishing between the different subtypes of ovarian cancer.

6.3 ROC–AUC and Classification Reliability

The model's sensitivity and specificity was evaluated using ROC analysis. As shown in Fig.Figure 4, the framework achieved near-perfect discrimination, with macro- and micro-average AUC values of 0.999. Class-specific AUCs ranged from 0.999 to 1.000, underscoring the model's ability to distinguish subtle histopathological variations across different subtypes.

Table 2: Classification report

Class	Precision	Recall	F1-Score	Support
Clear_Cell	0.9852	0.9852	0.9852	135
Endometrioid	0.9924	1.0000	0.9962	131
Mucinous	1.0000	0.9704	0.9850	135
Non_Cancerous	0.9635	0.9778	0.9706	135
Serous	0.9694	0.9794	0.9744	97
Accuracy		0.9826		633
Macro Avg	0.9821	0.9825	0.9823	633
Weighted Avg	0.9828	0.9826	0.9826	633

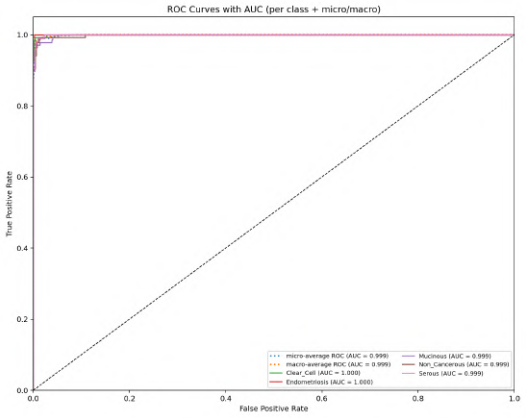


Fig. 4. ROC curves and AUC scores for each class.

7 Grad-CAM Visualization Analysis

To further investigate the interpretability[14] of the proposed CoAtNet-based framework, Gradient-weighted Class Activation Mapping (Grad-CAM)[12] was applied to visualize the spatial attention regions for each ovarian cancer subtype. Grad-CAM produces class-specific activation heatmaps by combining feature maps from the last convolutional layer with the corresponding gradients of the predicted class, thereby highlighting the regions that most influence the model's decision[13].

For each of the five classes, a representative histopathological image was selected to illustrate the interpretability of the model. As shown in Fig.Figure 5 through Fig.Figure 9, the visualizations are organized into three panels: (i) the original H&E-stained histopathology image, (ii) the Grad-CAM heatmap indicating areas of high model activation, and (iii) the superimposed overlay showing the correspondence between the heatmap and the tissue regions.

In the case of Clear Cell carcinoma Fig.Figure 5, the activation map strongly highlights clusters of tumor cells with clear cytoplasm and distinct boundaries, confirming that the model correctly attends to morphologically relevant patterns. For Endometrioid carcinoma Fig.Figure 6, the Grad-CAM intensity focuses on glandular and epithelial formations characteristic of endometrial-like morphology. The pattern of Grad-CAM activation as shown in Figure 7 depicts that the pattern of activation of mucinous carcinoma is highly localized to mucus-secreting epithelial cells and related gland cavities, thus illustrating the ability of the model to identify secretory histopathology.

As Figure 8 illustrates, the Grad-CAM activations of non-cancer tissue sample are located in the stromal components, which suggests the network is able to differentiate between normal and cancer tissue structures. Lastly, Figure 9 shows that the activation of the papillary structures with highly packed nuclei has very strong marks, which proves the accuracy of the model in detecting the presence of high-grade serous carcinoma features.These visualization results provide strong evidence of interpretability and transparency[15], essential for clinical trust and practical deployment in digital pathology.

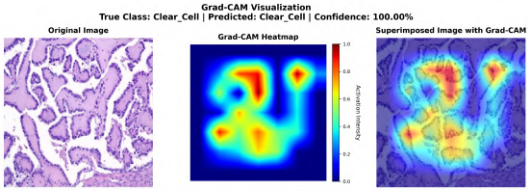


Fig. 5. Grad-CAM visualization for Clear Cell carcinoma

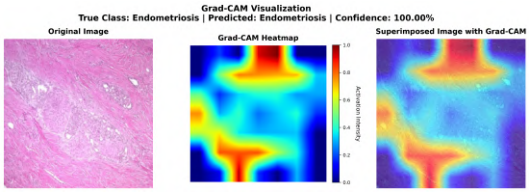


Fig. 6. Grad-CAM visualization for Endometrioid carcinoma

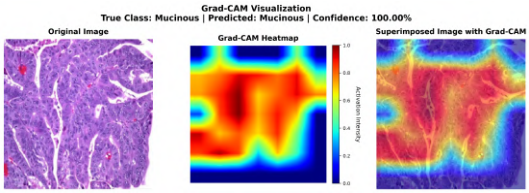


Fig. 7. Grad-CAM visualization for Mucinous carcinoma

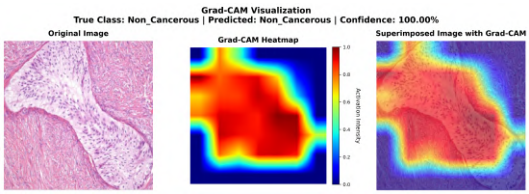


Fig. 8. Grad-CAM visualization for Non-cancerous ovarian tissue

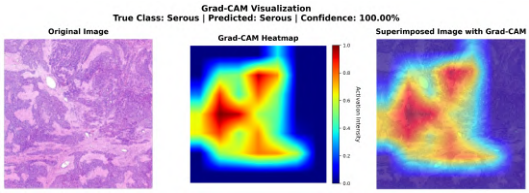


Fig. 9. Grad-CAM visualization for Serous carcinoma

8 Conclusion

This paper demonstrates a Convolution Attention model combined with Explainable Artificial Intelligence as an automated multi-classifier to classify the subtypes of ovarian cancer using histopathology images. Based on the CoAtNet architecture, the proposed model succeeds in combining convolutional feature extraction with the transformer-based attention mechanisms to identify both localized morphological features and the overall contextual features. The framework had a test accuracy of 98.26 per cent. and constantly high F1-scores among all five categories, namely, Clear Cell, Endometrioid, Mucinous, Serous, and Non-cancer tissues.

This addition of Grad-CAM-based visualizations helped to establish interpretability, as it would turn out the histopathologically relevant histonucles or epithelial border, stromal frameworks, and so on, and proved that the model predictions are informed by diagnostically relevant features. These outcomes of explainability[16] contribute to better clinical transparency and the confidence in the use of AI-supported pathology devices.

The results show that in digital histopathology explainable AI-supported hybrid convolution-attention models can achieve high accuracy and interpretability. These developments have the potential to support the introduction of deep learning models into the clinical process of diagnostics to allow the early and accurate detection of ovarian cancer.

References

[1] Radhakrishnan M, Sampathila N, Muralikrishna H, Swathi KS. Advancing ovarian cancer diagnosis through deep learning and explainable AI: a multiclassification approach. IEEE Access. 2024;12:116968–116986.

[2] Dai Z, Liu H, Le QV, Tan M. CoAtNet: marrying convolution and attention for all data sizes. Adv Neural Inf Process Syst. 2021.

[3] Xu W, Xu Y, Chang T, Tu Z. Co-scale conv-attentional image transformers. Proc IEEE Int Conf Comput Vis. 2021.

[4] Kasture KR, Choudhari D, Matte PN. Prediction and classification of ovarian cancer using enhanced deep convolutional neural network. Int J Eng Trends Technol. 2022;70(3).

[5] Lu M, Fan Z, Xu B, Chen L, Zheng X, Li J, et al. Using machine learning to predict ovarian cancer. Int J Med Inform. 2020.

[6] Zhang Y, Guo M, Wang L, Weng S, Xu H, Ren Y, et al. A tumor-infiltrating immune cells–related pseudogene signature based on machine learning predicts outcomes and immunotherapy responses in ovarian cancer. Cell Signal. 2023.

[7] Ayyoubzadeh SM, Ahmadi M, Yazdipour AB, Ghorbani-Bidkorpeh F, Ahmadi M. Prediction of ovarian cancer using artificial intelligence tools. Health Sci Rep. 2024.

[8] Hamidi F, et al. Identifying potential circulating miRNA biomarkers for the diagnosis and prediction of ovarian cancer using machine-learning approach: application of Boruta. Front Digit Health. 2023;5.

[9] Gao Y, et al. Deep learning-enabled pelvic ultrasound images for accurate diagnosis of ovarian cancer in China: a retrospective, multicentre, diagnostic study. Lancet Digit Health. 2022;4(3).

[10] Kwatra CV, Kaur H. A comprehensive investigation into the use of machine learning to forecast ovarian cancer. Proc ICCCNT. 2023.

[11] Liu Y, Lawson BC, Huang X, Broom BM, Weinstein JN. Prediction of ovarian cancer response to therapy based on deep learning analysis of histopathology images. Cancers (Basel). 2023;15(16).

[12] Chaddad A, Peng J, Xu J, Bouridane A. Survey of explainable AI techniques in healthcare. Sensors. 2023;23(2).

[13] Saeed W, Omlin C. Explainable AI (XAI): a systematic meta-survey of current challenges and future opportunities. Knowl Based Syst. 2023;263.

[14] Saraswat D, et al. Explainable AI for Healthcare 5.0: opportunities and challenges. IEEE Access. 2022;10.

[15] Bussmann N, Giudici P, Marinelli D, Papenbrock J. Explainable AI in fintech risk management. Front Artif Intell. 2020;3.

[16] Holzinger A. Explainable AI and multi-modal causability in medicine. i-com. 2021;19(3).

[17] Holzinger A, Langs G, Denk H, Zatloukal K, Müller H. Causability and explainability of artificial intelligence in medicine. Wiley Interdiscip Rev Data Min Knowl Discov. 2019;9(4).

[18] Ziyambe B, et al. A deep learning framework for the prediction and diagnosis of ovarian cancer in pre- and post-menopausal women. Diagnostics (Basel). 2023;13(10).

[19] Suganyadevi S, Seethalakshmi V, Balasamy K. A review on deep learning in medical image analysis. Int J Multimed Inf Retr. 2022;11(1).

[20] Akazawa M, Hashimoto K. Artificial intelligence in ovarian cancer diagnosis. Anticancer Res. 2020;40(8):4795–4800.