

Hybrid ResNet-ViT Framework for Endometrial Lesion Analysis: A Comparative Study of MRI and CT in Endometrial Cancer Classification

Omar H Abu-azzam¹, Amer Mahmoud Sindiani², Salem Alhatamleh³, Mohammad Amin³, Hamad Yahia Abu Mhanna⁴, Rola Madain², Hanan Fawaz Akhdar⁵, Hasan Gharaibeh⁶, Omar F Altaf², Eman Hussein Alshdaifat⁷, Tarfah Majed Alinad⁵, Fatimah Maashey⁵, Ahmad Nasayreh⁶, Ayah Bashkami⁸, Latifah Alghulayqah⁵

¹Department of Obstetrics and Gynecology, Faculty of Medicine, Mutah University, Al-Karak, Jordan; ²Department of Obstetrics and Gynecology, Faculty of Medicine, Jordan University of Science and Technology, Irbid, Jordan; ³Computer Science Department, Faculty of Information Technology and Computer Sciences, Yarmouk University, Irbid, Jordan; ⁴Department of Medical Imaging, Faculty of Allied Medical Sciences, Isra University, Amman, Jordan; ⁵Physics Department, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia; ⁶Artificial Intelligence and Data Innovation Office, King Hussein Cancer Center, Amman, Jordan; ⁷Department of Obstetrics and Gynecology, Faculty of Medicine, Yarmouk University, Irbid, Jordan; ⁸Department of Medical Laboratory Sciences, Al Balqa Applied University, Salt, Jordan

Correspondence: Hanan Fawaz Akhdar, Email hfakdar@imamu.edu.sa

Aim: This study aimed to evaluate and compare the diagnostic performance of computed tomography (CT) and magnetic resonance imaging (MRI) in the detection of endometrial cancer, using a deep learning approach.

Methods: Two endometrial image sets were obtained from King Abdullah University Hospital: the KAUH Endometrial Cancer MRI dataset (KAUH-ECMD) and the KAUH Endometrial Cancer CT dataset (KAUH-ECCTD), collected from 300 patients aged between 22 and 85. A hybrid deep learning model combining ResNet50 and Vision Transformer (ViT) was applied to classify the images into three categories: benign, malignant, and normal.

Results: The proposed ViTNet model achieved an accuracy of 90.24% in detecting endometrial cancer using MRI images and 86.99% using CT images. The MRI-based approach demonstrated superior diagnostic performance in detecting endometrial cancer compared to CT-based classification.

Conclusion: Deep learning models utilizing MRI and CT images demonstrate high accuracy in classifying endometrial cases. MRI in particular shows promise in supporting diagnostic workflows. Future work will focus on further validating the model's ability to evaluate depth of invasion and other prognostic features.

Keywords: endometrial neoplasms, magnetic resonance imaging, computed tomography, deep learning, radiomics

Introduction

The endometrium is the epithelial lining of the uterus, just like any cells in the body; they can be a source of malignant transformation. Endometrial cancer (EC) is the most common gynecologic malignancy in the US, causing more than 66,000 new cases and 13,000 deaths in 2023.¹ Most of these cases are endometrioid, also called type 1, which accounts for over 80% of uterine corpus cancers. Type 2 cancers include serous and clear cell carcinomas. They act more aggressively but fortunately make up only a smaller percentage of the total new cases.²

Molecularly, The Cancer Genome Atlas TCGA classifies endometrial cancer into four subtypes: DNA Polymerase Epsilon (POLE) ultra-mutated, microsatellite instability hypermutated, copy number low, and copy number high. The copy number indicates the changes that happen during cancer cell replication, which play an important role in determining the behavior of the neoplasm, with a higher copy number indicating a worse prognosis.³ Genetically, EC shares similarities with serous ovarian cancer, the basal-like subtype of breast cancer, and colorectal cancer. Serous tumors and some endometrioid tumors



often exhibit TP53 mutations, significant copy number alterations, and minimal methylation changes which normally cause the silencing of genes if it happen in the promoter region.⁴ In contrast, other endometrioid tumors tend to have fewer copy number alterations, rare Tumor Protein p53 (TP53) mutations, and frequent Phosphatase and Tensin Homolog (PTEN) and Kirsten Rat Sarcoma Viral Oncogene Homolog (KRAS) mutations.²

As a rule, in the endometrium, estrogen causes the proliferation of the epithelial lining, thereby increasing the chance of cancer development. In contrast, progesterone causes endometrial shedding, thus playing a protective factor against EC. Conditions that predominantly raise estrogen levels without raising progesterone can heighten the risk of EC, such as obesity, diabetes, and polycystic ovary syndrome.⁵ Hormone therapy that includes estrogen without progestin after menopause also raises this risk, as do rare ovarian tumors that release estrogen, such as granulosa cell tumors. A longer reproductive lifespan, due to early menstruation or late menopause, leads to greater estrogen exposure, further increasing risk. Additionally, women who have never been pregnant are more likely to develop EC. Aging is another factor, as the disease is most common after menopause. Obesity contributes to disrupting hormonal balance, with adipose tissue producing estrone in premenopausal women. Tamoxifen, a breast cancer treatment, carries a slight risk, though its benefits often outweigh the dangers.⁵

Due to the absence of standardized screening protocols for endometrial cancer, current recommendations emphasize symptom-driven evaluation rather than population-based screening. Diagnosis typically starts with a pelvic exam to check for irregularities in the reproductive organs, often alongside a Pap test, though Pap tests are not a reliable screening method for EC since the disease originates inside the uterus, and the Pap test collects cells via a brush or a spatula from the cervix, ideally from the transitional zone.⁶ However, if abnormal cells are detected in a Pap test, an endometrial biopsy may be required to confirm cancer by analyzing a tissue sample from the uterine lining. Other diagnostic procedures include dilation and curettage, which removes uterine tissue under anesthesia, and transvaginal ultrasound to detect abnormal endometrial thickening. If cancer is confirmed, staging tests such as hysterectomy, biopsies, chest X-rays, MRIs, or CT scans help determine if the cancer has spread beyond the uterus.⁷ Studies on non-coding RNAs, primarily Y RNAs,⁸ as well as miRNAs,⁹ underscore their implications in early diagnosis, prognosis, and fertility-preserving treatment of endometrial cancer. While our study focused on the use of images, it involved bodies of evidence comprising imaging data and liquid biopsy transcriptomic biomarkers (such as the proposed Endometrioid Endometrial RNA Index by¹⁰ to develop hybrid diagnostic pipelines that can guide personalized treatment.

In radiology, AI has demonstrated the potential to reduce radiologists' workload, minimize errors, and expedite diagnoses, particularly for conditions like breast, bone, liver, lung, and cardiovascular diseases.¹¹ AI-driven imaging, including CT, MRI, and ultrasound, enhances accuracy and reduces human error; however, human oversight remains necessary. While AI has shown promise in cancer detection, more research is needed to uncover its potential role in EC diagnosis and risk assessment and to come up with generalizable conclusions about the best way to implement it and integrate AI with EC radiology in a way to ensure patient safety and maintain confidentiality. Expanding AI applications in this area may enhance diagnostic workflows, particularly for underserved populations, but further research is needed to determine their long-term impact on outcomes.¹²

Because EC has been increasing in both incidence and mortality, particularly in the developed world, early detection and accurate diagnosis are crucial for improving patient outcomes. While artificial intelligence (AI) has been widely used in breast cancer imaging, its application in gynecologic cancers, including EC, remains underexplored.¹³ AI techniques, such as machine learning (ML) which extracts knowledge from data, and deep learning (DL) which is more complex and uses artificial neural networks to extract knowledge from less structured data, have the potential to enhance radiologic imaging by improving tumor classification, predicting metastasis, and aiding in staging.¹⁴ Traditional diagnostic imaging heavily relies on radiologists' expertise which can be variable depending on the radiologist's training, but AI can standardize and refine the process, making diagnosis more efficient. Machine learning algorithms analyze large datasets to distinguish between benign and malignant tumors, while convolutional neural networks (CNNs) mimic human brain functions to analyze vast amounts of data and can learn complex image features, improving accuracy in cancer detection. However, AI models require extensive training data to ensure reliability and avoid overfitting, in which a model is created that fits the training set so perfectly that it fails when applied to new sets of data. Radiomics, which is the extraction of quantitative data from medical images, helps make AI-driven analysis more computationally feasible, thus making the field of radiology most applicable to AI because images can be turned into lines of code.¹⁵

Preoperative assessment for endometrial cancer includes performing body scans, and we will be comparing both MRI and CT scans. MRI is the most accurate modality in assessing the endometrial disease involvement due to its preferred soft tissue contrast resolution.¹⁶ But it's the least favorable for patients with claustrophobia, as they tend to refuse the imaging or terminate it prematurely. CT scans are useful in the evaluation of more advanced disease with extrauterine spread, lymphadenopathy, and metastatic disease beyond the pelvis.⁶ Which is needed during preoperative assessment, as its results could alter the management plan.

This work aims to leverage deep learning advances to create reliable models for endometrial cancer detection. The main contributions of this study are as follows:

- This study presents two new datasets of CT and MRI scans of the endometrium obtained from King Abdullah University Hospital in Jordan to diagnose and detect endometrial cancer.
- A hybrid approach combining ResNet50 and ViT Transformer is proposed, and its performance is evaluated on two datasets collected from the hospital.
- A comparison was made between MRI and CT in the classification of endometrial cancer.
- A comprehensive analysis and evaluation of the proposed model was performed on the datasets and compared with a set of pre-trained models, and its performance was analyzed.

The following sections of this study are organized as follows: Section 2 provides a comprehensive review of the relevant literature in this field. Section 3 outlines the applied methodology, details the dataset, and provides a detailed explanation of the proposed model's structure. Section 4 presents the experimental results of the proposed model. Section 5 discusses the limitations and conclusions of the proposed model, and Section 6 presents conclusions and outlines future research prospects.

Related Work

MRI and CT imaging have facilitated the diagnosis and classification of endometrial tumors. In this regard, many strategies have been examined, including deep learning model-based approaches, radiomics-based approaches, and hybrid systems that categorize imagery from multiple modalities. While the studies demonstrate the advantages of CT in identifying endometrial cancer, they also highlight the limitations of the datasets, variations in imaging techniques, and the need for external validation. The contributions of various strategies to enhancing tumor classification and diagnostic accuracy will be elaborated upon below.

Recent advances in radiomics have demonstrated promising results in the early detection, molecular profiling, and staging of endometrial cancer. These techniques extract high-dimensional quantitative features from medical images, providing non-invasive biomarkers that complement traditional imaging and histopathology.¹⁷ Lymph node involvement remains a critical prognostic factor in endometrial cancer. While retroperitoneal lymphadenectomy has traditionally been used for staging, its therapeutic benefit remains controversial due to associated morbidity. Sentinel lymph node (SLN) mapping has emerged as a less invasive alternative with promising diagnostic accuracy and prognostic value. Recent evidence supports the use of SLN mapping for accurate nodal assessment while minimizing surgical risks.¹⁸ Incorporating SLN techniques alongside advanced imaging may optimize staging strategies and guide treatment decisions.

Machine Learning and Deep Learning Techniques in Endometrial Cancer Diagnosis

Currently, the methods used to anticipate the recurrence of endometrial cancer rely on costly tests and thorough analyses, making them difficult for many individuals to obtain. In¹⁹ the HECTOR model was introduced. It's a deep learning method that utilizes images of tissue samples and tumor stages to predict outcomes accurately. It achieved high scores on tests, indicating that it outperforms current methods and facilitates the creation of more personalized treatment plans. In,²⁰ researchers examined the use of urinary fluorescence spectroscopy combined with machine learning to diagnose endometrial cancer without invasive procedures. They got good results, with an accuracy of 79% and an Area Under the Receiver Operating Characteristic Curve (AUC) score of 90%. This method is a cheaper and more accurate option than regular invasive tests, which could help find problems earlier and make it easier to use in clinics. In²¹ presented EndoNet,

a model combining convolutional neural networks and vision transformers to classify endometrial cancer histology slides into high- and low-grade categories. Achieving an F1 score of 0.91 and AUC of 0.95 on internal tests and 0.86 for both metrics on external tests, EndoNet offers a promising, annotation-free tool for supporting pathologists in tumor grading.

According to,²² researchers discovered a unique protein pattern in cervicovaginal fluid that enables precise differentiation between those with endometrial cancer and those without. They achieved a high accuracy rate with a score of 0.95, meaning it correctly identified 91% of those with cancer and 86% of those without. These results show that there might be a simple and affordable way to check for endometrial cancer without surgery. More testing is needed before it can be used in hospitals. In²³ a group of blood tests has been found that can help diagnose endometrial cancer. This method had a high accuracy rate of about 82.8% to 83.1% and a measurement of 0.901 to 0.902. They used a technique called particle-enhanced laser desorption/ionization mass spectrometry along with machine learning. Testing these biomarkers can help us understand biology better and improve non-invasive methods for diagnosing EC. To quickly and accurately assess Microsatellite Instability (MSI) status in endometrial cancer using H&E-stained slides, a study in²⁴ presented a deep learning model. This model has very good results: it has an F-measure of 96%, accuracy of 94%, precision of 93%, and sensitivity of 100% for endometrioid carcinoma G1G2. The model works much better than the standard methods and is fast, taking only 1.03 seconds to analyze each slide. This shows it could be useful in real medical settings.

In,²⁵ researchers created a deep-learning model to forecast how patients with Atypical Endometrial Hyperplasia (AEH) and early-stage endometrial cancer (EC) would respond to hormone treatment. They used whole-slide images and achieved an accuracy score of 0.80. These results show that it is possible to create personalized treatment plans for AEH and early EC using predictions from images. In²⁶ suggested a method was suggested that combines different machine learning models, VGG16, DenseNet121, and MobileNetV2, to identify endometrial cancer in tissue images. This method reached an accuracy of 97.17%, an F-score of 96.79%, a sensitivity of 97.17%, and a precision of 97.25%. The results show that the model can help find endometrial cancer early and accurately.

In,²⁷ the researchers employed machine learning and explainable AI to effectively assess the age, tumor grade, invasion depth into the myometrium, and tumor size in cases of endometrial cancer. Their accuracy rates were very high: 98.8% for age, 98.6% for tumor grade, 95.1% for myometrial invasion, and 94.8% for tumor size. It also points out possible markers, especially the EWRS1 protein, related to these prediction factors. In²⁸ a way to diagnose uterine corpus endometrial cancer (UCEC) without surgery. This method uses small pieces of DNA found in the blood, and it showed great results: an area under the curve (AUC) of 0.991 during testing and 0.994 during validation, with a high accuracy of 95.5% for not mistakenly identifying healthy people and 98.5% for correctly identifying those who are sick. The model can greatly help find uterine cancer early, identifying 99% of stage I UCEC patients. This could increase the 5-year survival rate from 84% to 95%. To improve ultrasound images for spotting uterine tumors, in²⁹ used techniques to clean up the images, highlight different parts, and extract important information. This resulted in a 97.8% accuracy rate using a Random Forest classifier. The results show that using ultrasound with machine learning can successfully tell the difference between benign and cancerous uterine tumors. This approach can be a cheaper option than MRI.

In³⁰ developed a vision transformer-based model was developed for identifying endometrial cancer in histopathology images, accomplishing 99.36% precision, outflanking fine-tuning strategies like MobilenetV2, Xception, and VGG16. The study proposes that vision transformer models can essentially upgrade early location and move forward with understanding results in endometrial cancer. In³¹ investigated the utilization of Deep Learning (DL) models to foresee bungle repair Mismatch Repair (MMR) status in endometrial cancer utilizing H&E-stained slide pictures, accomplishing an AUROC of 0.91 with ResNet50. The proposed AI-based arrangement offers a cost-effective pre-screening device for atomic classification, lessening the requirement for more costly genetic tests. Whereas³² illustrated the possibility of utilizing Deep Learning (DL) to evaluate the degree of Myometrial Invasion (MI) in endometrial cancer from ultrasound pictures, with EfficientNet-B6 accomplishing an AUC of 0.814 and outflanking 15 radiologists in demonstrative precision. The model's predominant execution recommends its potential as an important device for clinical decision-making in EC treatment.

In³³ researchers recommended an AI-based approach for diagnosing uterine cancer. This method included separating tumor cells, training a deep learning model, and improving it using a technique called Pigeon-Inspired Optimization. They achieved a very high accuracy of 99.85% in five different categories. The results show that AI can help in finding and identifying uterine cancer earlier. In,³⁴ researchers created a 3D model called Prompt-nnUnet to automatically outline areas of concern and sensitive

organs in brachytherapy for endometrial cancer. Their model worked closely with expert radiologists and made the outlining process faster. The model showed great results with scores of 0.96 for heart region segmentation and 0.91 for rectum segmentation, making clinical work more efficient. In,³⁵ the authors introduced the ECMS-Net framework for separating endometrial cancer tumors in MRI images. This framework uses a classification model based on Transformers and a segmentation model called U2-Net. The method reached a classification accuracy of 98.5% and a maximum F1 score of 97% for tumor segmentation, showing great results for use in healthcare. Table 1 demonstrates a comparison of earlier endometrial cancer research.

Table 1 An Analysis of Earlier Research on Endometrial Cancer

Ref	Dataset	Approach	Model Used	Result
[19]	Tissue images, tumor stage data	Predict the recurrence of endometrial cancer	HECTOR (Deep Learning)	High accuracy and improved personalized treatment predictions.
[20]	Urine fluorescence data	Non-invasive diagnosis using spectroscopy	Machine Learning	Accuracy: 79%, AUC: 90%.
[21]	Histology slides	Tumor grading (high- vs low-grade)	EndoNet (CNN + Vision Transformers)	F1: 0.91, AUC: 0.95 (internal), F1 & AUC: 0.86 (external).
[22]	Cervico-vaginal fluid samples	Protein pattern analysis for cancer detection	Statistical Analysis	Accuracy: 91% (cancer), 86% (non-cancer).
[23]	Blood sample data	Biomarker-based non-invasive diagnosis	Mass Spectrometry + Machine Learning	Accuracy: ~82.8–83.1%, AUC: ~0.901–0.902.
[24]	H&E-stained slide images	Predict MSI status quickly and accurately	Deep Learning	F1: 96%, Accuracy: 94%, Sensitivity: 100%, Precision: 93%.
[25]	Whole-slide images	Predict hormone therapy response in AEH/early EC	Deep Learning	AUROC: 80%.
[26]	Tissue images	Cancer detection using ensemble ML models	VGG16, DenseNet121, MobileNetV2	Accuracy: 97.17%, F1: 96.79%, Sensitivity: 97.17%, Precision: 97.25%.
[27]	Clinical and histology data	Prediction of tumor metrics and marker detection	Machine Learning + Explainable AI	Accuracy: 94.8–98.8%. Identifies EWRS1 protein as a marker.
[28]	Circulating DNA samples	Non-invasive UCEC diagnosis	Machine Learning	AUC: 0.991 (testing), 0.994 (validation), Accuracy: 95.5%.
[29]	Ultrasound images	Tumor classification using image enhancement	Random Forest Classifier	Accuracy: 97.8%.
[30]	Histopathology images	Early detection using Transformers	Vision Transformers	Accuracy: 99.36%.
[31]	H&E-stained slide images	Predict MMR status	ResNet50	AUC: 91.00%
[32]	Ultrasound images	Assess myometrial invasion	EfficientNet-B6	AUC: 0.814. Outperforms radiologists.
[33]	Tumor cell datasets	Tumor cell classification using optimization	Deep Learning + Pigeon-Inspired Optimization	Accuracy: 99.85%.
[34]	MRI images	Segmentation for brachytherapy planning	Prompt-nnUnet	Segmentation Scores: 0.96 (heart), 0.91 (rectum).
[35]	MRI images	Tumor segmentation using a hybrid framework	ECMS-Net (Transformers + U2-Net)	Classification Accuracy: 98.5%, F1: 97.0%.

Other Datasets

Datasets used in endometrial cancer queries include a variety of clinical and imaging information. The Endometrial Cancer Dataset (2013–2024) provides histopathological and CT information, focusing on tumor types and stage,³⁶ while the Hormonal Variables and Endometrial Cancer Probability Dataset (2013) addresses risk factors such as balance and obesity in 173,519 women.³⁷ The International Endometrial Tumor Analysis (IETA) Ultrasound Dataset (2017) includes ultrasound images for tumor staging and evaluation in 1,538 cases,³⁸ and the Endometrial Cancer Imaging Dataset (2019) includes annual information on 61,880 cases, combining imaging procedures such as ultrasound, CT, and MRI for tumor assessment and staging.⁶ Table 2 provides a comparison of endometrial imaging datasets.

Methodology

This paper presents a comprehensive approach model for using advanced deep learning architecture to identify endometrial cancer. Data were obtained from King Abdullah University Hospital (KAUH-ECMD and KAUH-ECCTD), including CT and MRI images classified as normal, benign, and malignant. The preprocessing approach included image compression, normalization, dataset balancing, and segmentation to ensure high-quality data preparation. The proposed model ViTNet, which combines deep learning architecture and a hybrid model mixed with ResNet50 and a transformer-based vision model, is then used for easy feature extraction and representation learning. Low- and mid-level features are extracted from ResNet50, pre-trained on ImageNet, reconstructed into token sequences, and fed into Transformer Encoder using a multi-head self-attention mechanism for spatial dependency considerations. The encoded features are then subjected to global average pooling and fully connected layers that are passed to Softmax activation for classification. The training is performed in two steps: first, ResNet50 is frozen for transformer-based learning; Second, fine-tune the entire model to optimally extract and classify features as shown in Figure 1.

Data Collection from a Hospital (KAUH)

This study was approved by the Institutional Review Board (IRB) of King Abdullah University Hospital (KAUH), affiliated with the Jordan University of Science and Technology (JUST), Irbid, Jordan. The dataset includes 1084 computed tomography (CT) images and 1084 magnetic resonance imaging (MRI) images, all obtained from a shared cohort of 300 female patients aged between 22 and 85 years. Each patient contributed multiple image slices per modality, generating the total number of images. Imaging data were retrospectively collected from KAUH's radiology archive

Table 2 Summary of Endometrial Imaging Datasets

Dataset	Original Year	Number of Cases	Number of Images	Views	Image Type File	Patient Info	Dataset Type
Endometrial Cancer	2013–2024	Not specified	Not specified	Multi-view	Histopathology, CT	Limited	Clinical and research use
Hormonal Factors and Endometrial Cancer Risk	Aggregated data from 4 studies	2013	173,519 women (360 + 1378 cases)	N/A	N/A	Postmenopausal, Caucasian women	Prospective cohort data
IETA Sonographic Data	2017	1538	Not Specified	Grayscale, Doppler	Ultrasound	Median age 65 years, BMI 28.4, menopausal status, and symptoms	Prospective Multicenter
Endometrial Cancer Imaging	2019	Approx. 61,880 cases/year in the U.S.	Not specified	Transvaginal US, CT, MRI, PET/CT, PET/MRI	Ultrasound, CT, MRI, PET	Median age 68 years, includes both premenopausal and postmenopausal women	Retrospective/ Prospective

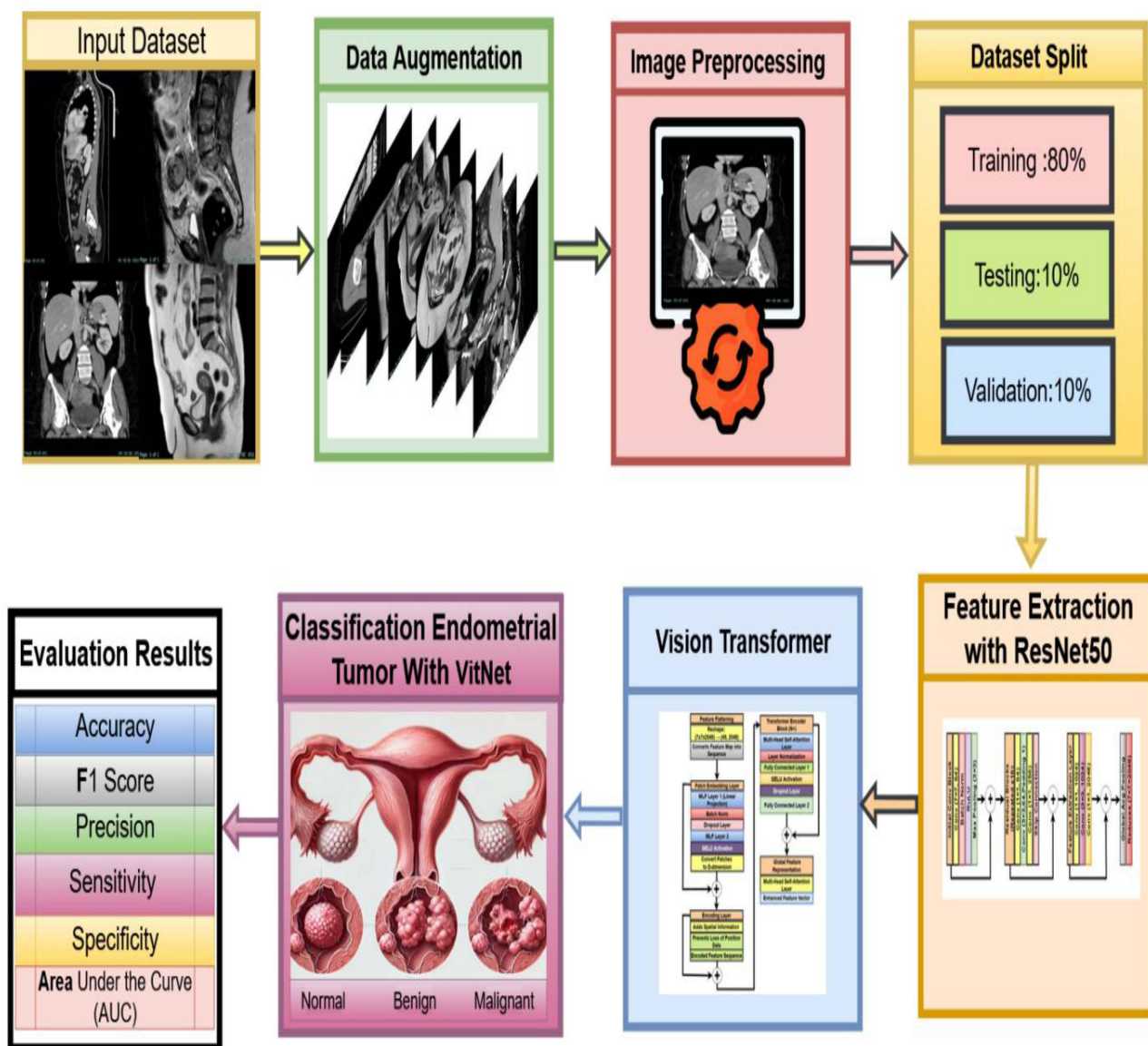


Figure 1 Process for methodology.

(PACS) between early 2019 and June 2024, while the data preparation process occurred between October 2024 and January 2025.

The final diagnostic classification for each case—normal, benign, or malignant—was based on the histopathological reports following biopsy or surgical intervention, which served as the reference standard for ground truth labeling. The categories were defined as follows:

- **Normal:** included images with physiological endometrial appearance, such as proliferative or secretory endometrium without pathology.
- **Benign:** included confirmed non-malignant abnormalities such as endometrial polyps or simple hyperplasia without atypia.
- **Malignant:** included only histologically confirmed endometrial cancer (EC) cases.

Inclusion criteria: Patients were included if they underwent either MRI or CT imaging of the pelvis at KAUH between 2019 and 2024 and had corresponding histopathological confirmation from biopsy or surgery.

Exclusion criteria: Patients with incomplete imaging studies, poor image quality (eg, due to motion artifacts), or missing histopathological confirmation were excluded from the dataset.

CT data included three standard anatomical views per patient: sagittal (lateral), coronal, and axial. For MRI, primarily T2-weighted sequences in axial, sagittal, and coronal planes were utilized for optimal soft-tissue contrast. All imaging data were anonymized before analysis. No personal identifiers were retained. Data were stored securely, in accordance with institutional protocols and the Declaration of Helsinki.

King Abdullah University Hospital Endometrial Cancer MRI Dataset (KAUH-ECMD)

The collection includes 1084 images taken by 300 women aged 22–85 years. Many scans were performed on these subjects, resulting in multiple image slices each time. The KAUH-ECMD dataset includes three MRI scans (sagittal, coronal, and axial) that were classified by hospital physicians as normal, benign, or malignant. Images are stored using MRI scans taken with an Ingenia Ambition 1.5T Sand machine and exported in JPG format with an average pixel resolution of 720×720. The distribution of the KAUH-ECMD dataset is presented in Table 3, with a sample from each group shown in Figure 2.

The King Abdullah University Hospital Endometrial Cancer CT Dataset (KAUH-ECCTD)

The collection includes 1084 photos taken by 300 women ranging in age from 22 to 85 years. Many of these people had repeated scans, each yielding many slices of images. The KAUH-ECCTD dataset comprises three segments of CT images (sagittal, coronal, and axial) that hospital clinicians evaluated as normal, benign, or malignant. Images are downloaded in JPG format with an average pixel resolution of 720×720 and saved using a Philips Brilliance 64-channel CT scanner. Table 4 illustrates the distribution of the KAUH-ECCTD dataset, whereas Figure 3 shows a sample from each group.

Data Augmentation and Preprocessing

Augmented data creates a more diverse training dataset by including variations of original images, thus improving the generalization of the models. This study uses augmentation techniques for endometrial MRI and CT images using the Image Data Generator.³⁹ The augmentations include random rotations, width, and height shifts, shearing, zooming in, and horizontal flipping. Such augmentations allow the model to learn how to recognize tumors while considering different angles, positions and sizes, thus reducing overfitting whilst improving robustness with unseen data. The dataset was divided into 80%, 10%, and 10% training, validation, and testing sets, respectively, using stratified sampling. This technique makes sure that the proportions of all classes appear in all subsets, which prevents imbalances in the classes. Stratified sampling ensures consistent distribution across dataset splits and augments the robustness of model training and evaluation due to representations of classes in the splits. Preprocessing input images for their standardization and compatibility with ResNet50.⁴⁰ The preprocessing function follows normalizing the pixel values by subtracting the mean and scaling based on ImageNet's distribution,⁴¹ as indicated in the equation below:

Table 3 The Quantity and Distribution of Photos in Each KAUH-ECMD Category

Case	Number of Images
Normal	292
Benign	385
Malignant	407
Total	1084

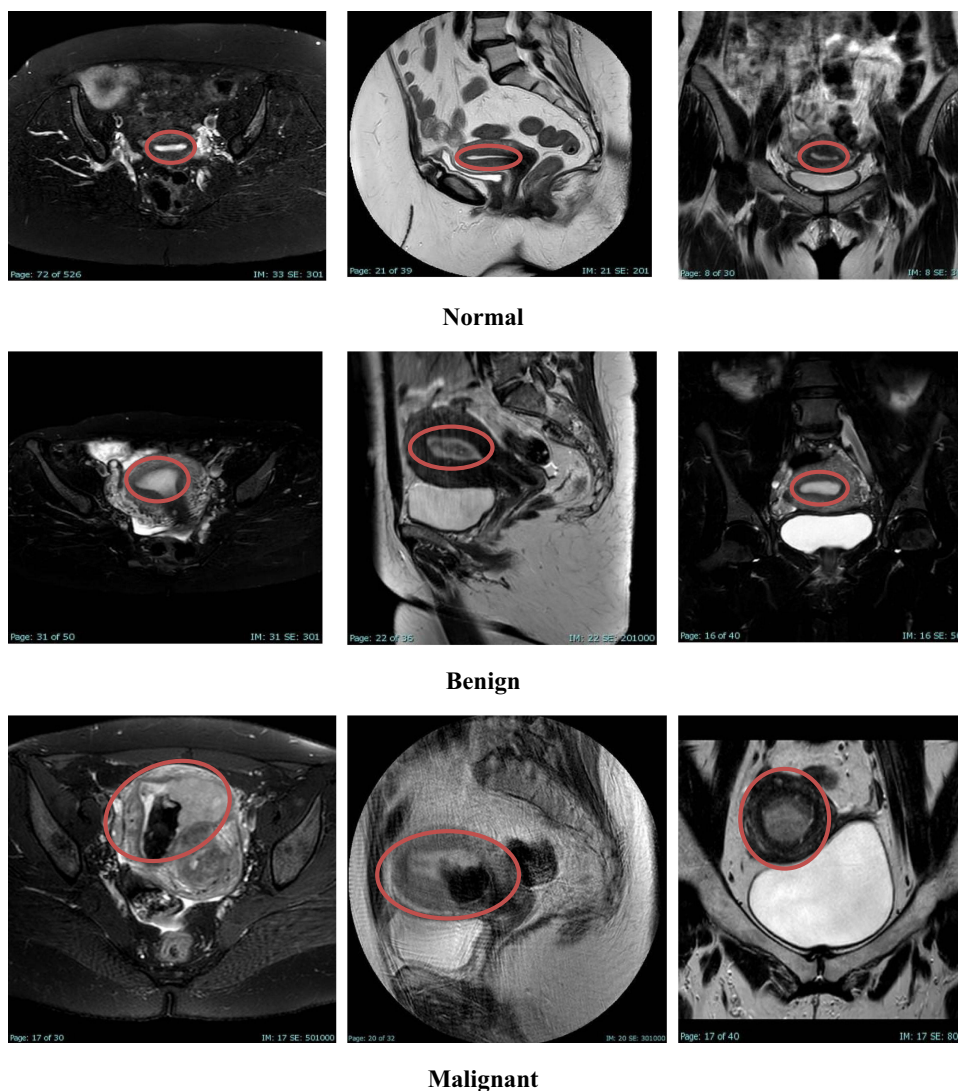


Figure 2 For example, images from the KAUH-ECMD dataset. The red circles indicate the location of the endometrium in normal cases (top row), benign lesions (middle row), and malignant lesions (bottom row).

$$x_{norm} = \frac{x - \mu}{\sigma} \tag{1}$$

The variables x are the pixel values and μ is the mean and σ is the corresponding standard deviation. Normalizing with such measures compensates for all variations due to light, contrast, or intensity for each image, ensuring that uniform

Table 4 The Quantity and Distribution of Photos in Each KAUH-ECCTD Category

Case	Number of Images
Normal	292
Benign	385
Malignant	407
Total	1084

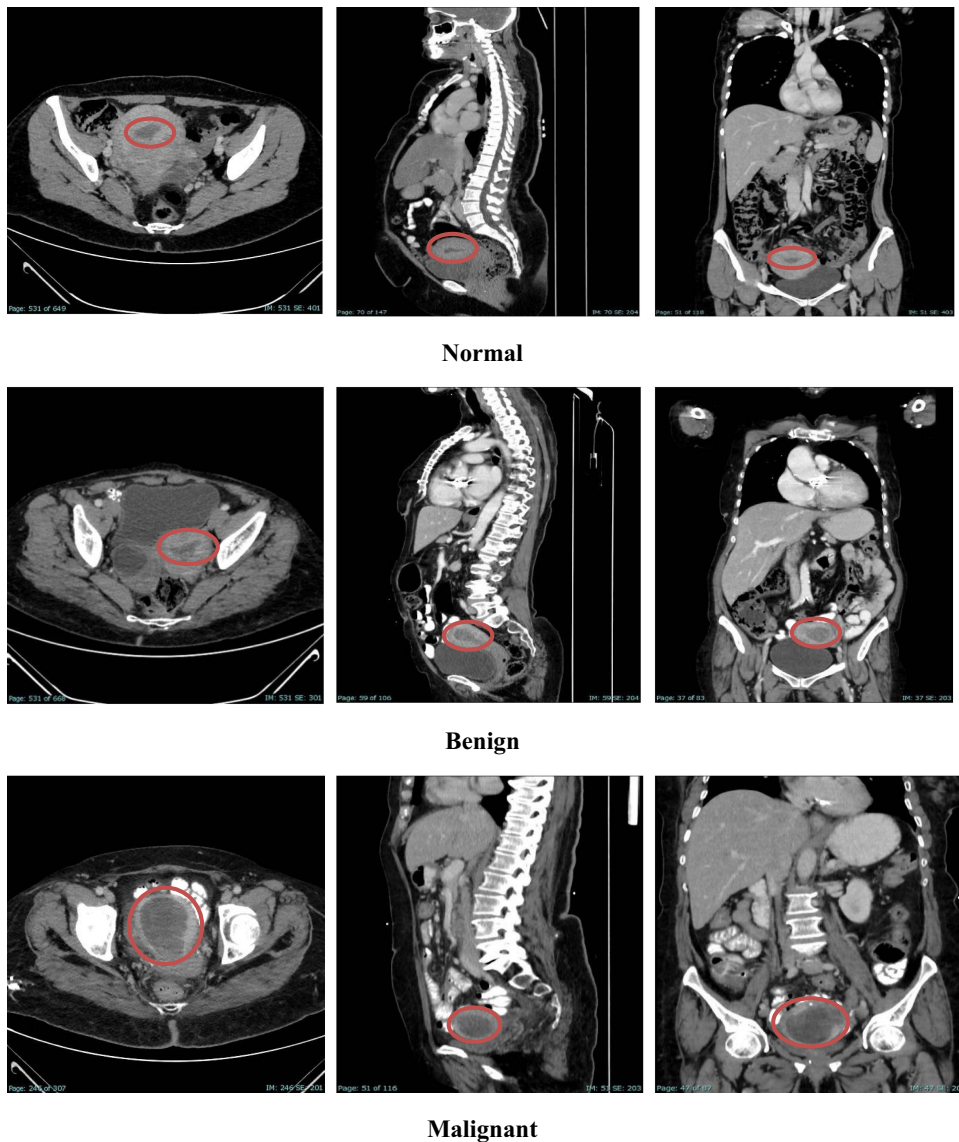


Figure 3 For example, images from the KAUH-ECCTD dataset. The red circles indicate the location of the endometrium in normal cases (top row), benign lesions (middle row), and malignant lesions (bottom row).

Red-Green-Blue (RGB) formatting is applied to all medical images.⁴² Accordingly, all images have been resized to 224×224 pixels to fit with the convolutional layers of ResNet50. By preprocessing the input data with ResNet50, we ensure that it becomes compatible with the feature space already learned by the model, subsequently improving transfer learning, extraction of relevant features, and enhancement of classification accuracy for endometrial tumor detection.

ResNet50 Model

ResNet50 is a deep convolutional neural network applied to overcome the vanishing gradient problem in extremely deep networks.⁴³ It was introduced together with the residual connections (or skip connections) idea, where the network can skip one or more layers and learn deeper networks with ease. This mechanism is used against the degradation problem, where deeper networks have lower performance as layers are piled on top of each other due to problems in gradient propagation. Residual block is the principal idea in ResNet50, and it contains a shortcut connection. This connection is mathematically represented by:

$$Output = F(x, \{W_i\}) + x \tag{2}$$

With $F(x, \{W_i\})$ representative of the convolutional layers, which in general, will include convolution, batch normalization,⁴⁴ and activation functions like ReLU, x will be the input to the residual block. The addition of the input x to the output of the convolutional layers creates a pathway for gradients to flow relatively easily through the residual connections during training, enhancing learning in deeper models. The other pivotal component in the ResNet50 is the activation function, generally the ReLU function:

$$ReLU(x) = \max(0, x) \tag{3}$$

There are non-linearities throughout convolutions due to ReLU thereby allowing the model to learn more complex patterns. Then, batch normalization is used to stabilize the learning by normalizing the output for each layer:

$$\hat{x} = \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}} \tag{4}$$

Where μ and σ^2 are the mean and variance of the activations respectively, and ϵ is a tiny constant to prevent division by zero. In this way, convergence speed and generalization can be improved. Figure 4 shows the structure of the ResNet50 architecture.

Vision Transformer (ViT) Model

This is transforming the vision transformer (ViT), which trains a transformer architecture layout for natural language processing to create its workings for image recognition tasks.⁴⁵ Affix a fixed-size patch upon an image with each flattened patch treated within as a token by the transformer. The power of self-attention makes it possible to capture distant patch relations with relevance focused directly into the area. Herein, the patch embedding process transforms each image patch into higher-dimensional space using linear projection:

$$z_i = Flatten(P_i)W \tag{5}$$

That P_i is the i^{th} image patch, flattening converts it into a 1D vector, while W is the learned projection matrix that projects the patch into a higher-dimensional space.⁴⁶ Positional encodings are appended to patch embeddings to provide spatial information regarding the positions of the patches since transformers do not capture spatial relationships by themselves:

$$z_i^{pos} = z_i + e_i \tag{6}$$

where e_i is the positional encoding for the i^{th} patch, which allows for the retention of spatial information during the processing.⁴⁷ Its central mechanism is self-attention, which determines attention scores between every pair of patches. The attention function is defined by:

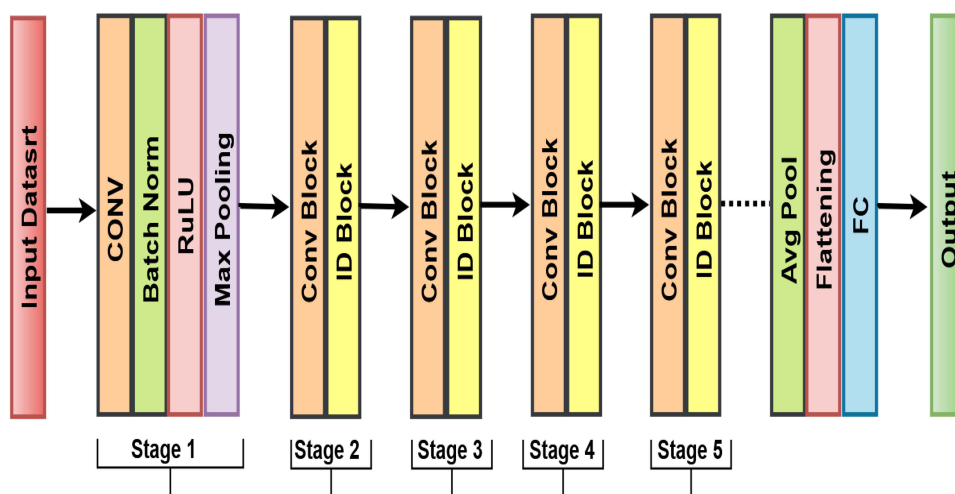


Figure 4 ResNet50 Methodology Architecture.

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

$Q, K,$ and V impulse represents the query key, and value matrices, respectively, while d_k is the dimension of the key vectors.⁴⁸ This mechanism allows the model to concentrate on different parts of the image via different contexts. The interaction can allow simultaneous capturing of multiple relationships and is called multi-head attention.

$$head_i = Attention\left(QW_i^Q, KW_i^{VK}, VW_i^V\right) \tag{8}$$

W^O denotes the output projection that is learned. With this, the model is enabled to capture its diversity in all different kinds of relationships concerning the input. Finally, the transformed tokens are fed through a feedforward network,⁴⁹ as follows:

$$FFN = max(0, zW_1 + b_1)W_2 + b_2 \tag{9}$$

are weight matrices and b_1, b_2 are biases. The operation is applied independently at each location in the sequence, allowing for further processing of information. Figure 5 shows the structure of Vision Transformer (ViT).

Proposed Model (VitNet)

ViTNet gets its name from a combination of ViT (Vision Transformer) and Net (Neural Network). It is hybrid. In contrast to classical deep learning models, which build on pure convolutional architectures, this one exploits the self-attention mechanisms of Transformers to be able to model global contextual information effectively. The name further reflects the modern trend in naming for deep-learning engineering practitioners, namely efficiencies, scalability, and novelty in medical imaging classification. ViTNet combines CNN-based feature extraction (ResNet50) with Transformer-based sequence modeling into a novel framework for endometrial MRI classification, balancing performance with interpretability and computational tractability. Figure 6 shows the structure of ViTNet.

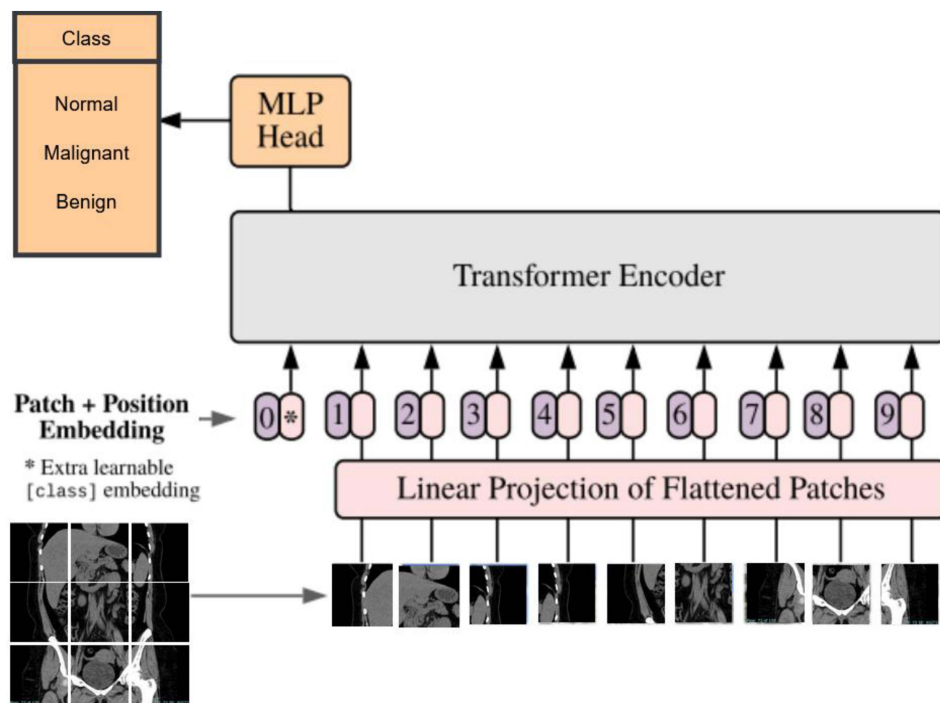


Figure 5 Flow chart for ViT Transformers.

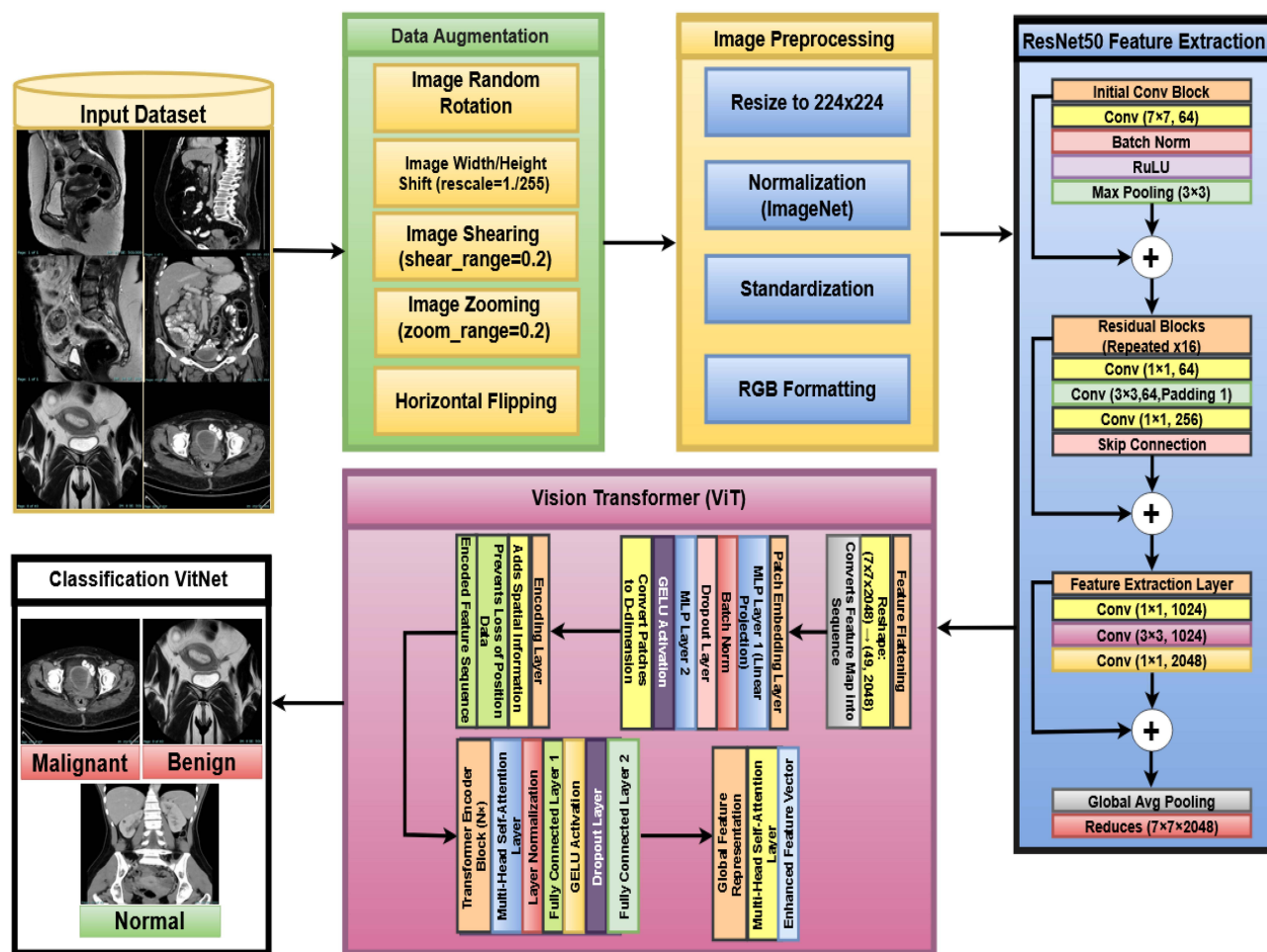


Figure 6 Structure of the proposed model ViTNet.

Feature Extraction Using ResNet50

The ResNet50 model, which has been pre-trained on the ImageNet dataset, is used as a feature extractor without the top layers.⁵⁰ ResNet50 is a deep convolutional neural network that is very helpful in overcoming the vanishing gradient problem by using residual connections. While the convolutional layers extract spatial features from MRI and CT images, they also maintain important structural information. In the present study, the fully connected (FC) layers are dropped, and only the convolutional base is used to extract meaningful feature representations. The images are then passed through the deep network for extracting spatial feature maps of size (7×7×2048). Thus, mathematically, a convolution layer is defined as follows:

$$Y(i,j) = \sum_m \sum_n X(i - m, j - n) \cdot K(m, n) \tag{10}$$

Here, $Y(i,j)$ is the output feature map at position (i,j) , X is the input image, K is the convolution kernel (filter), (m, n) denote kernel size. The residual connections support gradient flow in ResNet50:

$$Y = F(X) + X \tag{11}$$

Where $F(X)$ refers to the operation carried out by the convolution layers and Y is the output after the residual connection. Building upon these residual blocks, ResNet50 seeks to enhance feature learning without compromising deep networks.⁴³ The features that are eventually extracted (7 × 7×2048) are then fed to the subsequent transformer-based ViTNet Module, thus enhancing the classification of endometrial tumors into benign and malignant or normal cases. Figure 7 shows ResNet50 Feature Extraction Architecture.

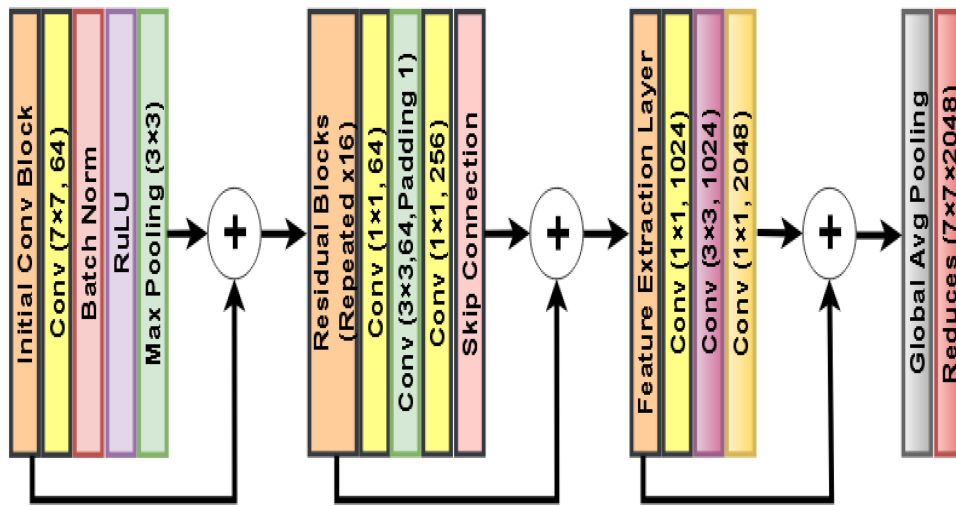


Figure 7 ResNet50 Feature Extraction Architecture.

Feature Transformation with Vision Transformer (ViT)

The feature map is flattened after extracting the spatial features out of ResNet50 and passed through a transformer encoder.⁵¹ This is an important step as it will enhance long-range dependencies and global relations among the features, something ResNet50 is not strong in handling. Patch embedding and positional encoding provide a different twist between the hybrid model working with feature map patches from the ResNet50 backbone compared to conventional Vision Transformers that directly use image patches.

ResNet50 gives a feature map of $7 \times 7 \times 2048$, which is reshaped into a sequence of 49 patches (each with 2048 features).⁵² Each patch is then embedded into a high-dimensional space. This embedding step is crucial for capturing richer representations of each feature. Apart from this, learned positional encoding is attached to each patch for spatial information, which is lost when flattening the input. This positional encoding gives the model a relative understanding of where those features reside in the input image. Formally, this can be expressed mathematically:

$$Z_0 = X + PE \tag{12}$$

Where X is the flattened sequence of ResNet50 features (ie, the reshaped 49 patches). PE is the positional embedding that is learned during training. Z_0 is therefore the resultant input to the transformer encoder endowed with contextual information.

Self-attention is applied multiple times in parallel (with different learned projections for each head) into a multi-head attention.⁵³ These attention heads are concatenated and projected to the final output. Thus, the final output for the attention mechanism is:

$$MHSA(X) = Concat(head_1, head_1, \dots, head_h) \tag{13}$$

Where $head_i$ is the output of the $i - th$ attention head. So, the model can focus on different parts of the feature map at once and be able to learn different aspects of the global relationships.

In one training episode, different parts of the feature map are sampled at once so that different aspects of the learning of the global relationships can take place. The self-attention output is followed by two operations: Layer Normalization and Feedforward Network (FFN).⁵⁴ Layer Normalization normalizes the output of the self-attention process, allowing the system to stabilize training. The key purpose of the operation is to normalize the input-to-next layer in terms of centering and scaling. Feedforward Network (FFN): After the output passes through the FFN, the straight-line functioning is broken, adding non-linearity to the model. An FFN usually consists of two linear transformation layers with a ReLU activation in between. Residual Connection During the backpropagation step, a residual connection is introduced to bypass the problem of gradient vanishing or exploding:

$$Y = FFN(LayerNorm(Z)) + Z \quad (14)$$

This ensures that models can store earlier representations while enhancing the gradient path during training. This has a rigorous process utilizing the ViT module, can easily capture global dependencies and long-range interactions among different feature patches. These abilities are fundamental to the enhancement of a hybrid model's performance during the classification of difficult medical images, such as endometrial MRIs and CT scans, for example. Figure 8 shows Feature Transformation with Vision Transformer (ViT).

Training Strategy VitNet

This training strategy consists of two stages for the proper extraction of features and deep integration of ResNet50 and ViT components.⁵⁵ These derivation stages aim to enable the learning of robust features at the initial phase before fine-tuning the model properly. Through this phase, exclusively training Transformer layers, the model can learn to concentrate on how to learn high-level representations by using ResNet50 frozen feature maps. All this time serves for stabilization actions with minimum and not drastic changes on the weights of the pretrained ResNet50 model.

This phase is where the Transformer structures were trained after which the model entered fine-tuning during which the weights of both the ResNet50 backbone and Vision Transformer were optimized together. Unlocking the last 10 layers of ResNet50 allows the model to update weights during training for performance improvement of the feature extraction mechanism to be tailored to the actual task-fused actualization of the overall classification accuracy.

Joint Optimization: Parameters of ResNet50 and Transformer are updated in parallel.⁵⁶ The fine-tuning phase enables simultaneous adjustment of both components so that the features from ResNet50 harmonize well with the long-range attention mechanism of the Transformer. This phase becomes very critical because it allows the model to adjust the ResNet50 features to the peculiarities of the dataset while tuning the Transformer attention mechanism for the best performance.

The Adam Optimizer is used for this setup because of the adaptive learning rate and dealing with sparse gradients. Further, the learning rate is set to $(\alpha = 1 \times 10^{-5})$, a small enough value to prevent heavy weight updates during this

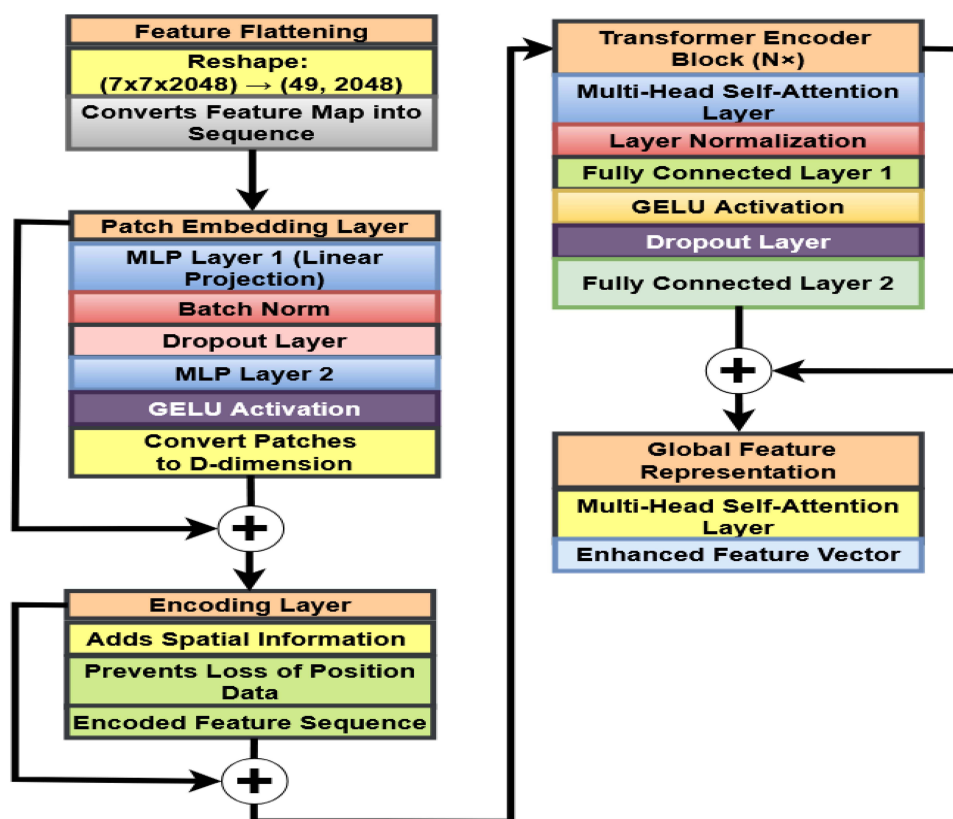


Figure 8 Feature Transformation Architecture Using Vision Transformer (ViT).

phase, enforcing a rather smooth convergence of the model during training. Categorical cross-entropy is used as the loss function to be optimized since the model is performing a classification task (ie, endometrial MRI classification). Categorical cross-entropy is suitable for multi-class classification problems, guiding the model to minimize the distance between its predicted probability distribution and the true label. The categorical cross-entropy loss can be defined as:

$$Loss = - \sum_{i=1}^N y_i \log(\hat{y}_i) \tag{15}$$

Where N is the number of classes, y_i is the true label for the class i , \hat{y}_i is the predicted probability for the class i . Table 5 shows the hyperparameter configurations used to simulate the proposed model using the Python implementation.

Classification with VitNet

All dimensions are pooled at the final output of the transformer encoder; this is an important step that reduces the whole dimension of the data while preserving the critical properties needed to perform classification. In endometrial tumor classification, particularly in this case, MRI and CT. GAP generally applies to the output of the Transformer encoder, which collects spatial features based on encoded representations by averaging their values across their spatial dimensions.⁴² It is useful for most such tasks, most specifically, medical imaging tasks such as the classification of tumors occurring in the endometrial, as it reduces the dimensionality of this feature map while retaining information along key spatial axes. The dimensionality-reduced version better generalizes across the different types of MRI and CT images that contain complex tumor growth patterns with surrounding tissue structures. Mathematically, GAP computes the average value for each feature map across the spatial dimensions:

$$GAP(X) = \frac{1}{H \times W} \sum_{l=1}^H \sum_{j=1}^W X_{i,j} \tag{16}$$

Converting high-dimensional outputs of the Transformer encoder into a more concise way of presenting it to the next layers is one more technique. After the application of GAP, the reduced data features go through a fully connected (FC) layer.⁵⁷ Acting as a classifier, this layer maps the output from GAP into the desired target number of classes, Benign, Malignant, and Normal endometrial tumor, the final output. The output pooled features from the fully connected layer undergo a linear transformation to enable the model to differentiate between different tumor classification types. When it comes to endometrial tumor classification, this layer captures and converts meaningful information from complex MRI and CT data of different types

Table 5 Values of the Hyperparameters for Training the Suggested and Other Models

Hyperparameters	Value
Image size	224 × 224
Batch size	16
Learning rate	1e-5
Optimizer	Adam
Activation function	ReLU
Units for Attention Module	256
Activation Function	ReLU
Units in Dense Layer	512
L2 Regularization Strength	0.01
Dropout Rate	0.5
Initial Learning Rate	1e-5

of tumors into categories by producing intelligible constructs through the feature pooling process in this layer. Ultimately, the output from the FC layer comes down to logits, which is the raw predictions of the model.

The logits are then passed through a Softmax activation function after the FC layer.⁵⁸ The logits are normalized through the Softmax function and rendered into a probability distribution over these classes such that all the probabilities add up to 1. This step is very critical for classification since the final layer output gives the probability of each class and enables the model to make an informed selection regarding the classification of the input data. The function is defined as:

$$P(y = c|x) = \frac{e^{z_c}}{\sum_j e^{z_j}} \quad (17)$$

Where $P(y = c|x)$ indicates the probability of the input belonging to the class c , z_c is the logit value meant for class c , z_j for each class j , and the denominator is the summation of the exponentials over all classes.

Endometrial tumors can be classified as benign, malignant, or normal by means of Softmax output within, which gives the model a probabilistic perspective on whether a tumor falls under any of those classifications. This proves useful in medical imaging, a domain where classification can make all the difference in terms of accuracy and reliability to diagnose patients correctly.

Techniques for Transfer Learning

These are neural network-based models pre-trained on large databases and have the capacity of capturing great feature diversity from input data.⁵⁹ These serve as the basis for a wide variety of applications, from image classification to the translation of languages and the detection of objects. Tuning through optimization, or the fine-tuning of these models, is an excellent method of modeling a retraining process to attain efficiency and accuracy for a specific application. Therefore, these pre-trained models can be efficiently and correctly applied in practice. In addition to the good performance with small datasets, this transfer learning setup also accelerates computation time and saves computer resources. In this study, we evaluate the proposed model VitNet for the classification of endometrial tumor MRI and CT images against the established ones including MobileNetV2, VGG16, VGG19, ResNet50, and ViT Transformer. Selection of the models was made in terms of their varied architectures and established success across a wide range of image classification tasks. Thus, by studying these models' performance concerning medical imaging, we hope to find a fit model for endometrial tumor classification. As shown in Figure 9.

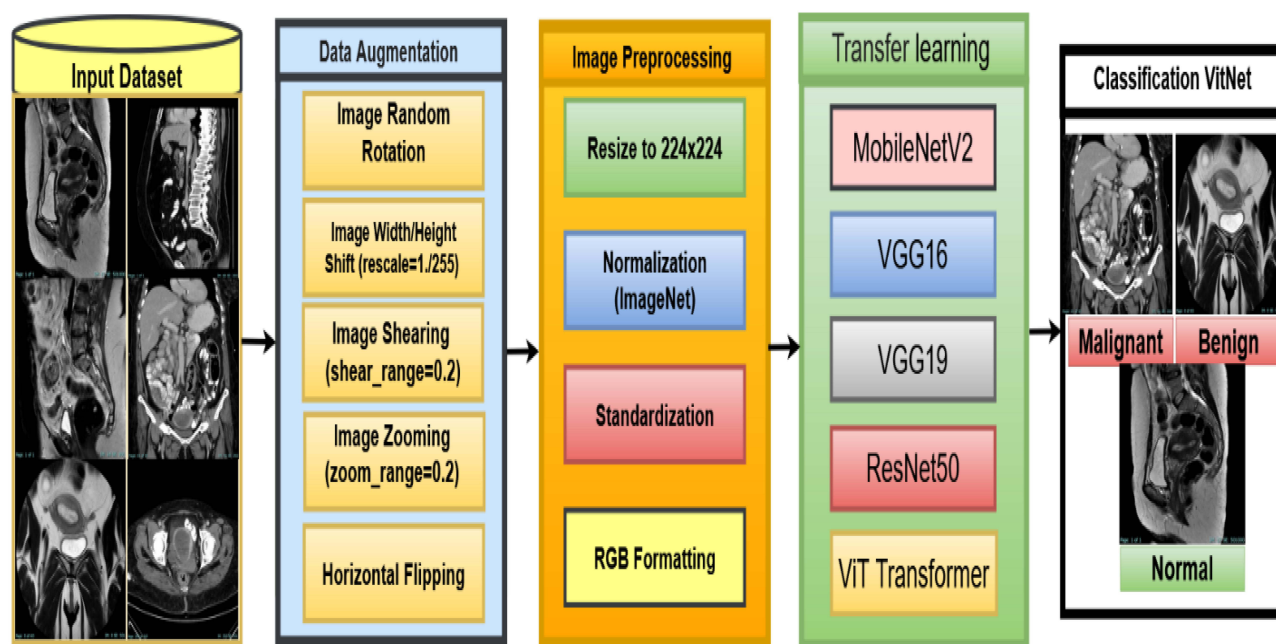


Figure 9 Working Methods of Transfer Learning.

Performance Evaluation Model

Through the investigation of these metrics, we will examine the ability of the model to diagnose endometrial tumors from MRI and CT images. Sensitivity checks for missed cases while specificity ensures that normal cases do not get misclassified, and precision minimizes unnecessary interventions. These three indicators must be balanced to develop a credible AI-based tool for endometrial diagnosis. This would be a very important assessment to carry out to determine the effectiveness of the model in diagnosing normal and benign as well as malignant endometrial tumors.

The accuracy of the model is defined as the overall correctness of its classifications across all tumor categories. TP (True Positive): Correctly classified malignant or benign tumor cases. TN (True Negative): Correctly classified normal cases. FP (False Positive): Normal cases misclassified as tumors. FN (False Negative): Tumor cases misclassified as normal.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (18)$$

Precision or Positive Predictive Value (PPV) tells the overall proportion of labeled benign or malignant cases from the predicted tumor cases. Thus, high precision ensures that very few healthy patients would be classified as tumor patients and benefit from unnecessary medical procedures.

$$Precision = \frac{TP}{TP + FP} \quad (19)$$

Sensitivity (Recall) or True Positive Rate (TPR) indicates the ability of the model to detect as many tumor cases as possible without leaving out any. Therefore, high sensitivity is more of a requirement in medicine because missing out on the presence of a tumor (false negative) can postpone the treatment.

$$Sensitivity = \frac{TP}{TP + FN} \quad (20)$$

Specificity or True Negative Rate (TNR) indicates how well the model identifies normal cases. In this way, a high specificity reduces the false positives so that normal cases are flagged as tumors unnecessarily.

$$Specificity = \frac{TN}{TN + FP} \quad (21)$$

F1 score balances precision and sensitivity; thus, it is well suited to the cases of imbalanced datasets when one out of two classes is less frequent (for example, malignant tumors). The high F1 score indicates that the model correctly classifies both tumor and normal cases without favoring one over the other.

$$F_1 - Score = 2 \frac{Precision \times Sensitivity}{Precision + Sensitivity} \quad (22)$$

Area Under the Curve (AUC - ROC) was used to evaluate the model's ability to discriminate between the three classes: normal, benign, and malignant. The ROC (Receiver Operating Characteristic) curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various classification thresholds. The AUC represents the probability that the model ranks a randomly chosen positive instance higher than a randomly chosen negative one. A higher AUC indicates better overall classification performance. The AUC is calculated as the integral of the TPR over the FPR:

$$AUC = \int_0^1 TPR d(FPR) \quad (23)$$

The False Positive Rate (FPR) is defined as the proportion of normal cases incorrectly classified as benign or malignant. In ROC plots, the diagonal line represents random classification, while curves closer to the upper-left corner indicate stronger discriminative performance.

Confusion matrix: A confusion matrix was constructed to evaluate the classification performance of the model by comparing predicted labels with actual ground truth labels. It provides a detailed breakdown of true positives, false positives, true negatives, and false negatives for each class. In our multiclass setting (normal, benign, and malignant), the confusion matrix shows how many cases from each actual category were correctly classified or misclassified into another.

This tool enables a more nuanced understanding of model behavior beyond overall accuracy, and it serves as the foundation for calculating performance metrics such as precision, recall (sensitivity), F1-score, and accuracy. Percentage values were included alongside raw counts in each matrix cell to provide normalized comparisons, and a color gradient was applied to visually represent classification strength and error distribution.

Computing Environment

All trials used Visual Studio Code (VS Code) on Windows 11 Pro. The training and evaluation support used an Intel Core i7-12700K processor with 16 GB of RAM and an NVIDIA RTX 4060 Ti-X GPU with 8 GB of memory. Model training took an average of approximately 5.8 minutes for 50 training sessions. However, each training session took approximately 7 seconds. GPU utilization was consistently above 90% during training. These facts show that the model is quite efficient when it comes to computing and can be used for technical purposes.

Result Analysis

This work compares MRI and CT in classifying endometrial cancer by developing a hybrid model, ViTNet, which integrates ResNet50 and Vision Transformer (ViT) architectures to categorize endometrial MRI and CT images into three classes: benign, malignant, and normal. The KAUH-ECMD and KAUH-ECCTD datasets were acquired from King Abdullah University Hospital in Jordan. The two datasets were partitioned into training (80%), validation (10%), and test (10%) subsets.

Model Performance Evaluation and Analysis on KAUH-ECMD

This section calculates each model's accuracy, precision, sensitivity, specificity, F1 score, and AUC. The KAUH-ECMD dataset is used to assess the ViTNet model's performance, which is then compared to five other models: MobileNetV2, VGG16, VGG19, ResNet50, and ViT Transformer. Table 6 shows how the models classified endometrial MRI scans as normal, benign, or malignant. With an accuracy of 90.24%, the suggested ViTNet model exceeds all previous models. The results reveal an F1 score of 90.21%, 90.63% precision, 90.63% sensitivity, and 95.09% specificity. The accuracy of the ViT Transformer and ResNet50 models is 70.73% and 53.36%, respectively. This demonstrates successfully the suggested methodology that detects endometrial cancer in MRI scans. With an accuracy of 87.80% and a specificity of 93.74%, the MobileNetV2 model finished second in the picture classification test, outperforming other models and demonstrating its ability to quickly extract and categorize complicated information from photos. Figure 10 shows a graphical illustration of the model's effectiveness.

Figure 11 displays the confusion matrices for various deep learning models employed to classify MRI data about endometrial cancer. Confusion matrices evaluate expected and actual classifications across three categories (normal, malignant, and benign) to measure model efficacy. The ViTNet model operates effectively because of its high accuracy and minimal misclassification. ViTNet maintains an equitable distribution across all classes, whereas VGG19 and ResNet50 encounter difficulties with the Normal class. MobileNetV2 exhibits comparable performance to earlier models; however, it demonstrates a slightly higher rate of misclassification. The Normal class adversely affects VGG19, resulting

Table 6 Model Performance with the KAUH-ECMD

Model	Accuracy	Precision	Sensitivity	Specificity	F1 Score
MobileNetV2	87.80	87.86	87.80	93.74	87.80
VGG16	81.30	83.21	81.30	90.25	81.35
VGG19	67.48	74.51	67.48	80.96	65.90
ResNet50	53.36	56.02	53.65	75.42	53.11
ViT Transformer	70.73	71.74	70.73	84.23	70.35
ViTNet	90.24	90.63	90.24	95.09	90.21

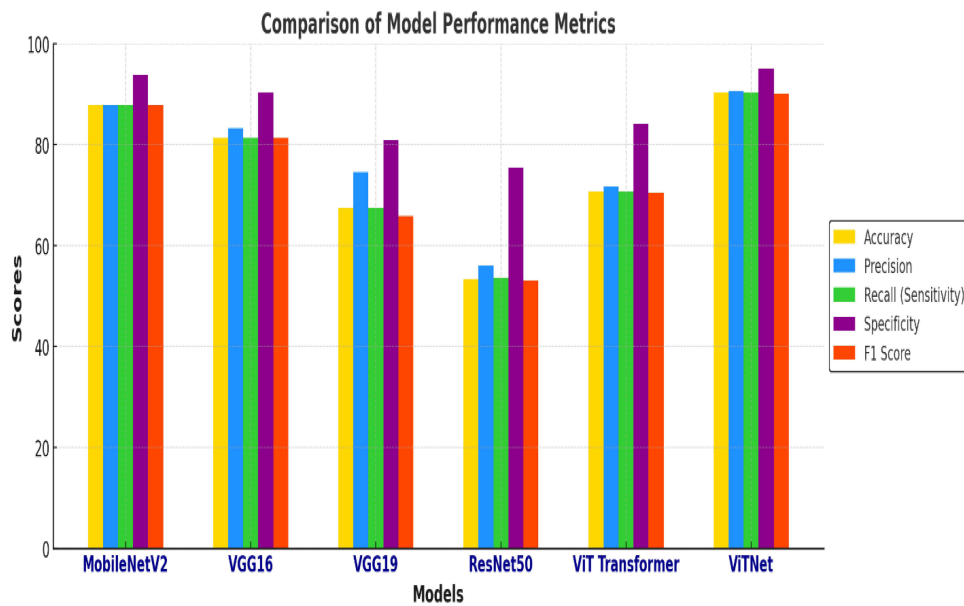


Figure 10 Performance comparison of models on the KAUH-ECMD dataset.

in increased misclassification. ResNet50 generates a greater number of misclassifications, particularly within the Normal and Malignant categories. Despite its excellent performance, the ViT Transformer exhibits a higher misclassification rate than ViTNet.

Model Performance Evaluation and Analysis on KAUH-ECCTD

In this section, we analyze the performance of the proposed ViTNet model on the KAUH-ECCTD dataset and compare it to five other models: ResNet50, InceptionV3, NASNetMobile, MobileNetV2, and VGG19. Table 7 demonstrates the models' ability to classify CT images as normal, benign, or malignant. The suggested ViTNet model has an accuracy of 86.99%, exceeding all previous models. It also has a precision of 87.51%, sensitivity of 86.99%, and specificity of 93.42%. This indicates the efficacy of the proposed technique for detecting endometrial cancer in CT images. The VGG16 model ranks second in the image classification test, with an accuracy of 85.36% and a specificity of 92.51%, surpassing other models and demonstrating its ability to efficiently extract and categorize complex information from images. MobileNetV2 and VGG19 were highly accurate in picture categorization, with 82.11% and 82.92%, respectively. In classifying endometrial cancer CT scans, ViT Transformer and ResNet50 performed the least effectively among all models; ViT Transformer achieved an accuracy of 75.60%, while ResNet50 achieved the lowest accuracy of 72.35%, demonstrating the effectiveness of combining ViT Transformer and ResNet50 into a single model. Figure 12 shows a graphical illustration of the model's effectiveness.

Figure 13 shows the confusion matrices of six different deep learning models - MobileNetV2, VGG16, VGG19, ResNet50, ViT Transformer, and ViTNet - used to classify endometrial CT scans into three classes: benign, malignant, and normal. The results show that ViTNet and VGG19 show superior performance, as evidenced by the high diagonal values. ResNet50 and ViT Transformer show a higher frequency of misclassifications, especially within the malignant and normal classes. MobileNetV2 and VGG16 show similar performance, albeit with some inaccuracies, especially in distinguishing between benign and typical cases.

Comparison of MRI and CT in Endometrial Cancer Classification Using ViTNet

In this study, a single hybrid deep learning model (ResNet50 combined with ViT) was employed for both the MRI KAUH-ECMD and the CT KAUH-ECCTD datasets. The purpose was to compare the diagnostic performance of the two imaging modalities using the same AI framework, rather than comparing different AI architectures. Based on an analysis of the KAUH-ECMD and KAUH-ECCTD datasets, the ViTNet model achieves better diagnostic performance using MRI

Confusion Matrices for Different Models

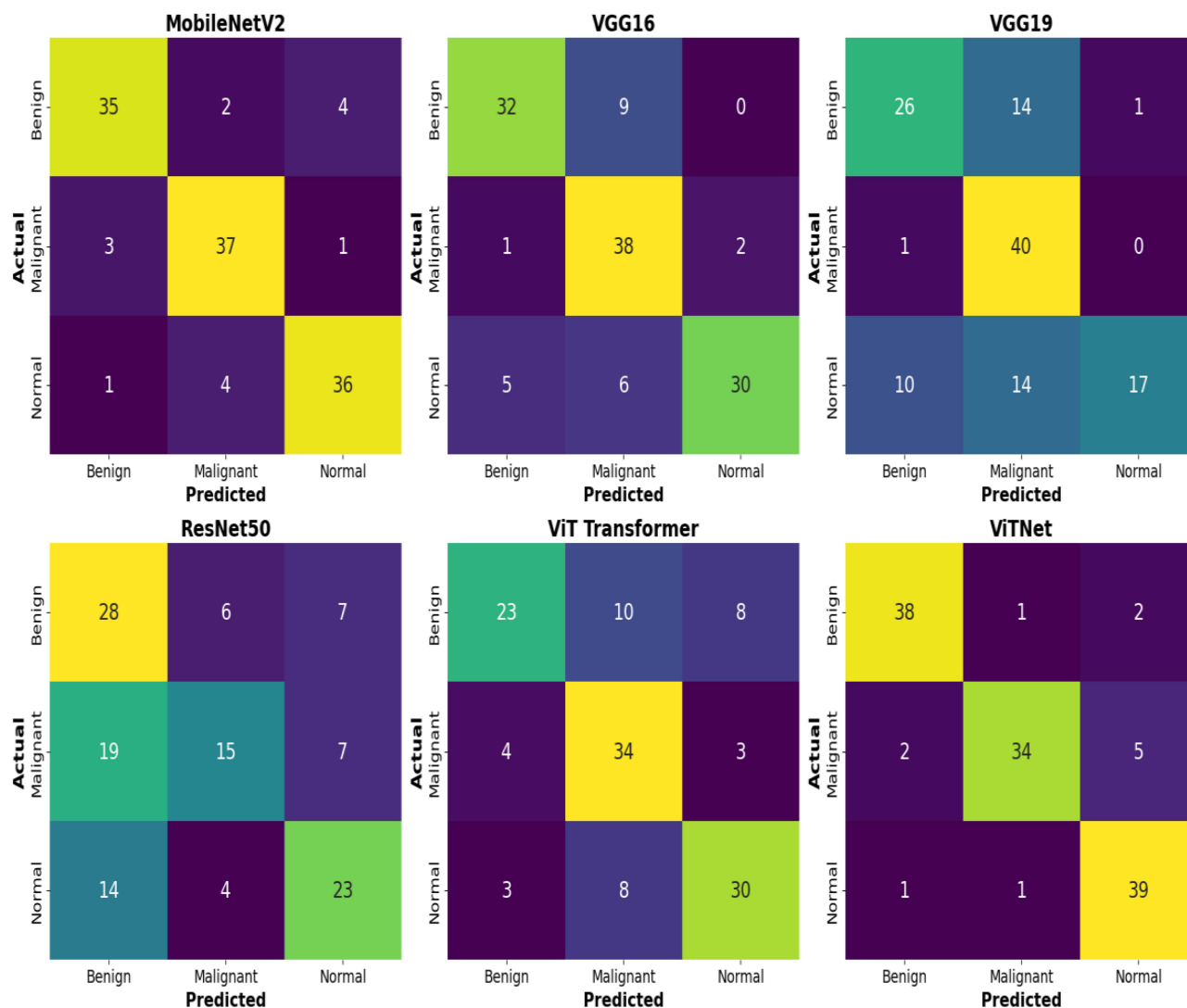


Figure 11 KAUH-ECMD Confusion Matrices.

than CT, with an accuracy of 90.24% and an area under the curve of 97.49%. The better the model performs, the closer the curves are to the upper left corner. The ROC curves that plot each class’s true positive rate (TPR) against false positive rate (FPR) are displayed in Figure 14. Strong classification performance is indicated by the AUC (area under the

Table 7 Model Performance with the KAUH-ECCTD

Model	Accuracy	Precision	Sensitivity	Specificity	F1 Score
MobileNetV2	82.11	82.51	82.11	90.85	82.00
VGG16	85.36	86.49	85.36	92.51	85.39
VGG19	82.92	84.68	82.92	91.41	82.84
ResNet50	72.35	74.27	72.35	85.87	72.13
ViT Transformer	75.60	78.77	75.60	87.57	75.45
ViTNet	86.99	87.51	86.99	93.42	87.06

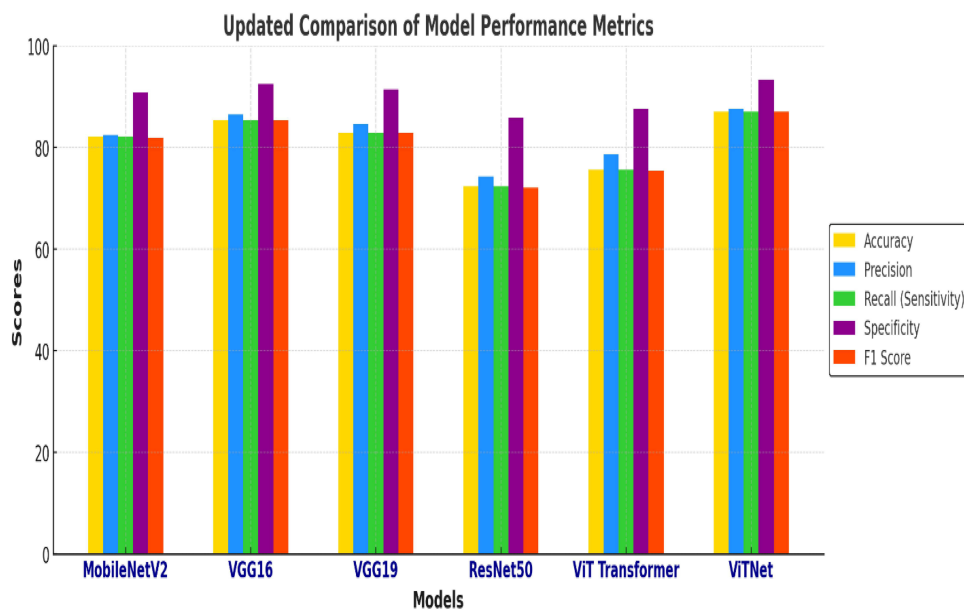


Figure 12 Performance comparison of models on the KAUH-ECCTD dataset.

curve) values, which show that both benign and malignant conditions achieve 0.97 and 0.98, respectively. A random classifier is represented by the dotted diagonal line; the better the model does, the closer the curves are to the upper left corner.

The model achieved an accuracy of 86.99% and an AUC of 95.89% in CT imaging diagnosis. The model achieved AUC values of 0.95 for benign and malignant cases and 0.98 for normal cases, as depicted in Figure 15. The model's discriminatory capability enhances as these curves approach the upper left corner; a random estimate is depicted by the dashed diagonal line.

Although MRI demonstrated superior diagnostic performance compared to CT, its higher cost and limited availability, especially in resource-constrained settings, may restrict its routine use. CT, being more widely available and cost-effective, may therefore remain a practical imaging modality in many healthcare systems, with MRI serving as a complementary tool where resources allow.

Discussion

In the study, a novel hybrid deep learning method has been developed mixing ResNet50 and Vision Transformer (ViT) architectures to classify endometrial lesions through MRI and CT imaging. The emphasis was laid on intending to evaluate and compare the diagnostic potential of two imaging modalities through a strong hybrid model which encapsulates ResNet's capability of spatial feature extraction and ViT's global contextual attention. The importance of this domain lies in its potential to enhance the early detection of endometrial cancer, which is crucial for effective treatment strategies and prognosis. The proposed model shows advanced capability in discriminating MRI data with an accuracy of 90.24%, thus highlighting how important it is to integrate transformer architectures into clinical diagnostic workflows. A hybrid architecture enhanced diagnostic accuracy and provided a more scalable and reproducible framework for medical image analysis in oncology, with significant implications for both clinical practice and future research in AI-driven cancer diagnostics.

The ViTNet framework for the classification of endometrial lesions is technically correct and is considered to have the potential for converging with evolving clinical paradigms in precision oncology. The latest publications of⁶⁰ emphasize the role of AI in gynecologic oncology for accurate risk stratification, treatment response prediction, and tailored patient care. More precisely, our model would help improve diagnostic accuracy, thereby preventing undue delays, invasive procedures in cases where these are not necessary, and the resulting occurrences of reduced clinical efficiency, affecting patient safety.

Confusion Matrices for Different Models

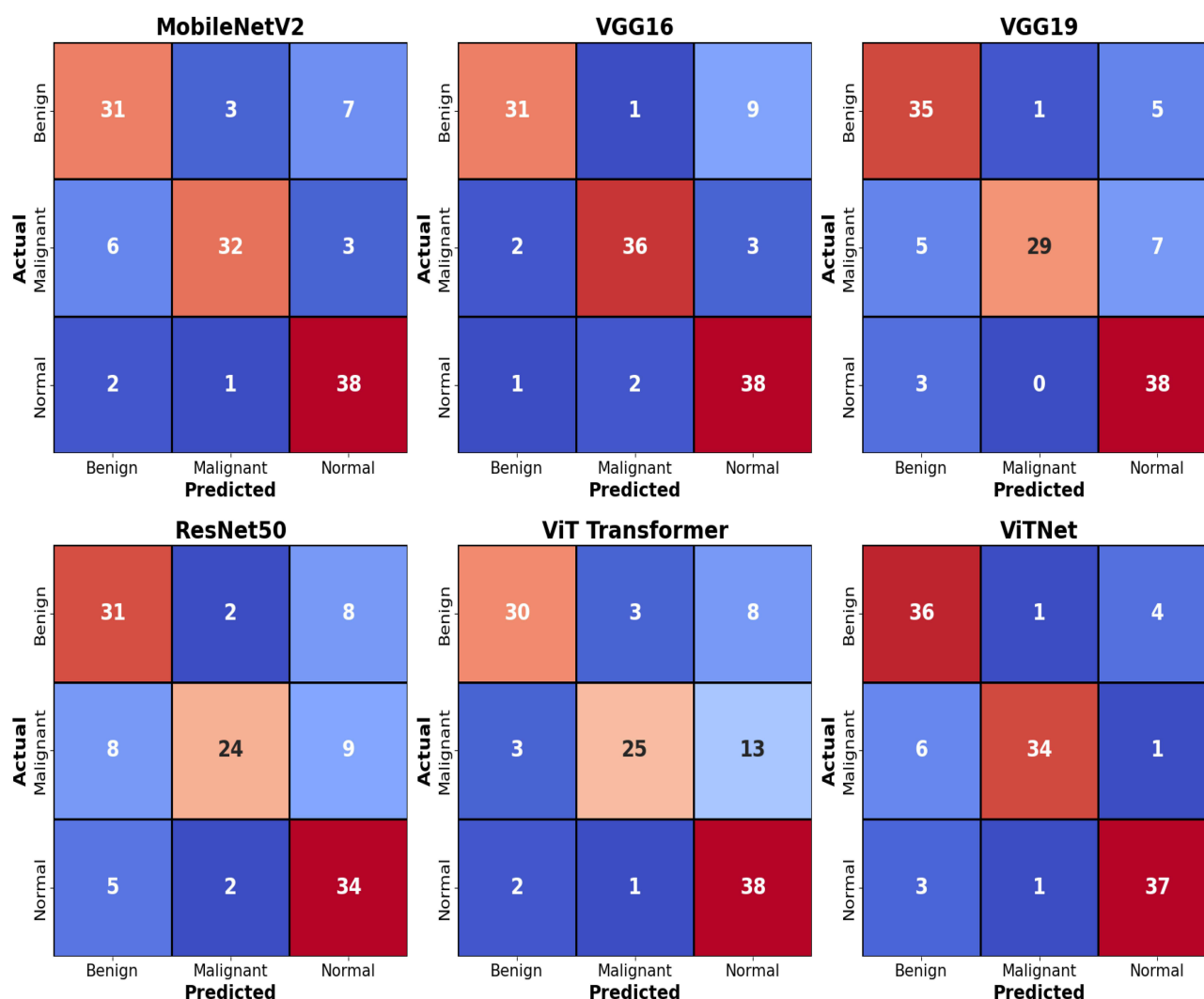


Figure 13 KAUH-ECCTD Confusion Matrices.

Preoperative risk stratification is demonstrated with strong potential through MRI-based radiomics and AI-integrated imaging, a demonstration carried out by.⁶¹ Yue et al⁶² developed a clinical-radiomics DL model based on MRI to classify endometrial cancer molecular subtypes, reporting macro-AUCs of 0.79 (internal) and 0.74 (external). Also, Qi et al⁶³ developed an AI-assisted MRI-based model (with attention mechanisms) for risk stratification and recurrence prediction, with AUCs >0.91 in several settings. In a more comparable CT-based work, Coada et al⁶⁴ showed that CT radiomic features can predict risk of recurrence with test AUCs between 0.86 and 0.90. These studies support our results, especially regarding the strong performance of MRI, and highlight that CT has potential too, though often slightly inferior in diagnostic accuracy. The research proposes⁶⁵ a hybrid model based on MobileNetV2 with dual attention and particle swarm optimization (PSO) mechanisms to accurately and efficiently classify endometrial cancer CT images. The model demonstrated superiority over traditional models, achieving an accuracy of 86.07%, and has the potential to be adopted as a supportive tool for early diagnosis and treatment planning.

The superior performance of MRI over CT in our study is consistent with prior literature and may be attributed to MRI's superior soft-tissue contrast, which allows clearer visualization of endometrial morphology and tumor boundaries. T2-weighted MRI sequences in particular provide enhanced delineation of the endometrium, myometrium, and potential

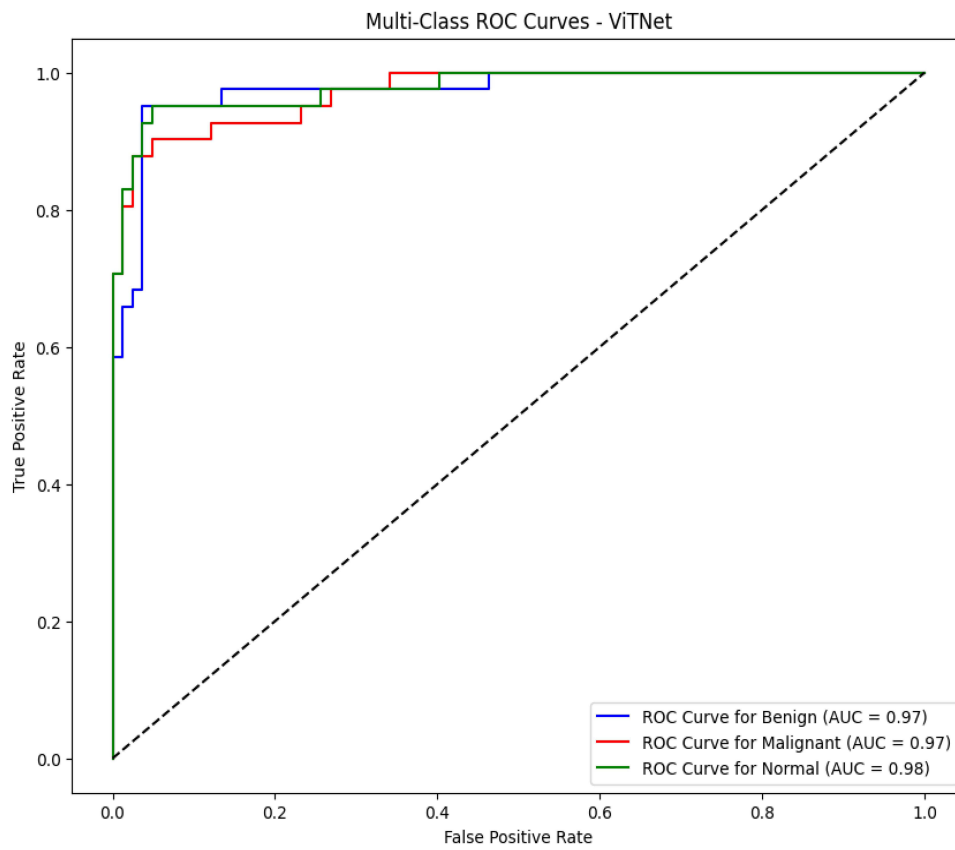


Figure 14 ROC curves for the classification of endometrial cancer using KAUH-ECMD.

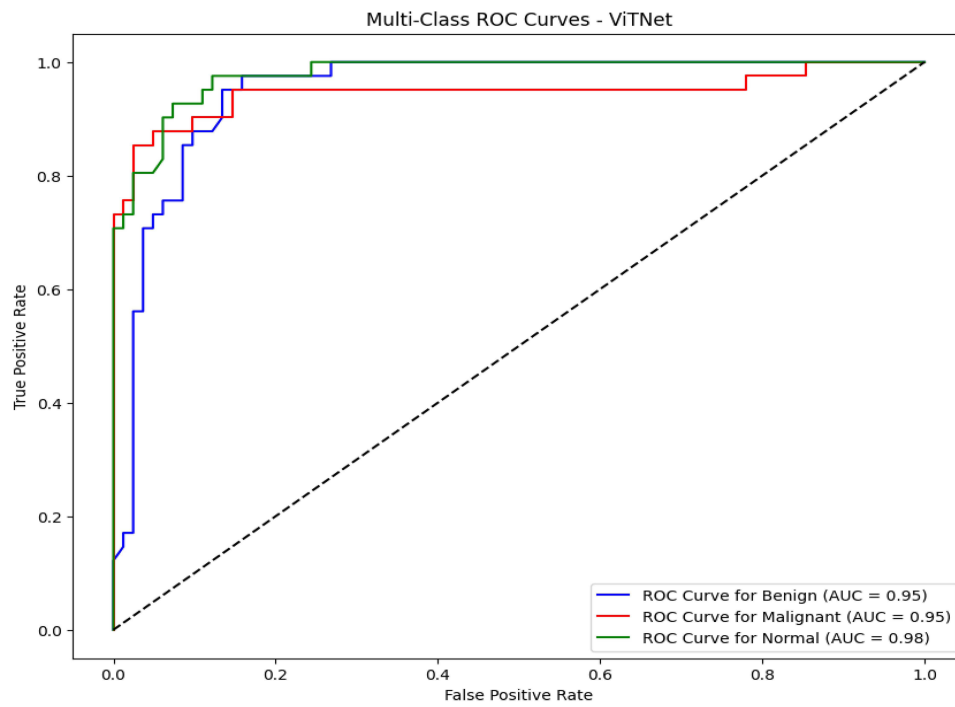


Figure 15 ROC curves for the classification of endometrial cancer using KAUH-ECCTD.

tumor invasion, features that are less well characterized on CT due to its lower tissue contrast resolution. These qualitative advantages of MRI likely enabled the deep learning model to extract more informative features, resulting in higher classification accuracy. Future work may further incorporate radiologist-guided qualitative analyses to better understand and interpret modality-specific differences.” Although MRI demonstrated superior diagnostic performance compared to CT, its higher cost and limited availability, especially in resource-constrained settings, may restrict its routine use. CT, being more widely available and cost-effective, may therefore remain a practical imaging modality in many healthcare systems, with MRI serving as a complementary tool where resources allow.

Our model uses deep learning and, in its present form, could be used in the extraction of radiomic features related to histopathological or molecular subtypes in the construction of multiparametric diagnostic models. Due to recent advancements in precision oncology, treatment tailoring is increasingly being considered with a side wash of image-based phenotyping in parallel with genomic and transcriptomic profiling. Our hybrid ViTNet framework paves the way for considering multimodal approaches, where radiological imaging signatures from MRI and CT are combined with molecular subtypes, either defined by TCGA or by the expression profiles of specific ncRNAs. Emerging lncRNAs, miRNAs, and circRNAs in tumor initiation, progression, and resistance to therapy provide a very compelling opportunity to non-invasively predict molecular alterations via informatics-aided advanced imaging analytics. Similar deep learning frameworks have been applied in breast, ovarian, and colorectal cancer stratification and thus could serve as an extension or adaptation for other neoplastic conditions. Such an integration is a game changer for diagnostic workflows, paving the way for the optimization of individual.

Limitations of the Study

This study has several limitations that should be acknowledged. First, the dataset was collected from a single center, which may limit the generalizability of the model to other institutions with different imaging protocols or equipment. This single-center dependency raises concerns about external validity, as the model’s performance might not be directly transferable to broader clinical practice without additional validation. Second, although the dataset size was reasonable, class imbalance, particularly between normal, benign, and malignant cases, may have affected the model’s learning dynamics. Third, the retrospective nature of data collection could introduce selection bias, as imaging was performed based on clinical suspicion rather than standardized screening. In addition, while deep learning models showed promising performance, their interpretability remains a challenge. Although we incorporated some model explainability techniques, the decision-making process of complex architectures such as transformers is still not fully transparent. Future work should explore more interpretable AI methods and validate the model across multi-center, prospective datasets with heterogeneous scanners and imaging protocols to enhance generalizability. In addition, the retrospective collection of data based on clinical suspicion may introduce selection bias, as patients without symptoms or those undergoing routine screening were not represented. As a result, the dataset may be enriched for more clinically apparent cases, potentially inflating the observed model performance compared to what might be expected in a general screening population.

Conclusion and Future Works

This study evaluated the performance of deep learning techniques in classifying endometrial MRI and CT images into normal, benign, and malignant categories. Using two novel datasets (KAUH-ECMD for MRI and KAUH-ECCTD for CT), we developed a hybrid model combining ResNet50 and Vision Transformer (ViT). The model achieved an accuracy of 90.24% on MRI data and 86.99% on CT data, demonstrating its potential to support radiological assessment of endometrial tumors.

While this study did not directly evaluate myometrial invasion or other staging features, our future work aims to expand the model’s capabilities to include such assessments. Additionally, we plan to develop a Computer-Aided Diagnosis (CAD) system that integrates multimodal imaging to assist clinicians in the early detection and classification of endometrial abnormalities. To improve representativeness and clinical reliability, we intend to conduct collaborative prospective studies across at least three tertiary hospitals, with larger and more diverse datasets, to further validate the model’s performance and applicability.

Data and Code Availability Statements

Data and code will be provided upon request to mohammadamin5111999@gmail.com.

Ethical Approval

This study was approved by the Institutional Review Board (IRB) of Jordan University of Science and Technology (JUST)/King Abdullah University Hospital (KAUH) (Approval No. 21/171/2024). The requirement for informed patient consent was specifically waived by the IRB, as the study is a retrospective analysis involving only de-identified electronic medical records. All data were handled in accordance with relevant ethical guidelines, and patient confidentiality was strictly maintained. This study was conducted in compliance with the ethical standards of the institutional research committee, Helsinki Declaration.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU), (grant number IMSIU-DDRSP2501).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Zaim SNM, Nafi SNM, Zawawi N, Adnan WFW. Clinical and histopathological evaluation of patients with endometrial cancer in a University Hospital: seven-year experience. *Malays J Pathol.* 2024;46(3):413–421.
2. Akhtar M, Hyassat SA, Elaiwy O, Rashid S, Al-Nabet ADMH. Classification of endometrial carcinoma: new perspectives beyond morphology. *Adv Anat Pathol.* 2019;26(6):421–427. doi:10.1097/PAP.0000000000000251
3. Dugo E, Piva F, Giulietti M, Giannella L, Ciavattini A, Gough L. Copy number variations in endometrial cancer: from biological significance to clinical utility. *Int J Gynecological Cancer.* 2024;34(7):1089–1097. doi:10.1136/ijgc-2024-005295
4. Lakshminarasimhan R, Liang G. The role of DNA methylation in cancer. *DNA Methyltransferases-Role and Function.* 2016;151–172.
5. Raglan O, Kalliala I, Markozannes G, et al. Risk factors for endometrial cancer: an umbrella review of the literature. *Int, J, Cancer.* 2019;145(7):1719–1730. doi:10.1002/ijc.31961
6. Faria SC, Devine CE, Rao B, Sagebiel T, Bhosale P. Imaging and staging of endometrial cancer. In: *Seminars in Ultrasound, CT and MRI.* Elsevier; 2019:287–294.
7. Sbarra M, Lupinelli M, Brook OR, Venkatesan AM, Nougaret S. Imaging of endometrial cancer. *Radiologic Clinics.* 2023;61(4):609–625. doi:10.1016/j.rcl.2023.02.007
8. Lefebvre TL, Ueno Y, Dohan A, et al. Development and validation of multiparametric MRI-based radiomics models for preoperative risk stratification of endometrial cancer. *Radiology.* 2022;305(2):375–386. doi:10.1148/radiol.212873
9. Piergentili R, Gullo G, Basile G, et al. Circulating miRNAs as a tool for early diagnosis of endometrial cancer—implications for the fertility-sparing process: clinical, biological, and legal aspects. *Int J Mol Sci.* 2023;24(14):11356. doi:10.3390/ijms241411356
10. Nief CA, Hammer PM, Wang A, et al. Endometrioid endometrial RNA index predicts recurrence in stage I patients. *Clin Cancer Res.* 2024;30(13):2801–2811. doi:10.1158/1078-0432.CCR-23-3158
11. Najjar R. Redefining radiology: a review of artificial intelligence integration in medical imaging. *Diagnostics.* 2023;13(17):2760. doi:10.3390/diagnostics13172760
12. Butt SR, Soulat A, Lal PM, et al. Impact of artificial intelligence on the diagnosis, treatment and prognosis of endometrial cancer. *Ann Med Surg.* 2024;86(3):1531–1539. doi:10.1097/MS9.0000000000001733
13. Shrestha P, Poudyal B, Yadollahi S, et al. A systematic review on the use of artificial intelligence in gynecologic imaging—background, state of the art, and future directions. *Gynecol Oncol.* 2022;166(3):596–605. doi:10.1016/j.ygyno.2022.07.024
14. Xiong L, Chen C, Lin Y, Mao W, Song Z. A computer-aided determining method for the myometrial infiltration depth of early endometrial cancer on MRI images. *Biomed Eng Online.* 2023;22(1):103. doi:10.1186/s12938-023-01169-w
15. Van Timmeren JE, Cester D, Tanadini-Lang S, Alkadhi H, Baessler B. Radiomics in medical imaging—'how-to' guide and critical reflection. *Insights Imaging.* 2020;11(1):91. doi:10.1186/s13244-020-00887-2
16. Munn Z, Moola S, Lisy K, Riitano D, Murphy F. Claustrophobia in magnetic resonance imaging: a systematic review and meta-analysis. *Radiography.* 2015;21(2):e59–e63. doi:10.1016/j.radi.2014.12.004

17. Di Donato V, Kontopantelis E, Cuccu I, et al. Magnetic resonance imaging-radiomics in endometrial cancer: a systematic review and meta-analysis. *Int J Gynecological Cancer*. 2023;33(7):1070–1076. doi:10.1136/ijgc-2023-004313
18. Cuccu I, et al. Sentinel node mapping in high-intermediate and high-risk endometrial cancer: analysis of 5-year oncologic outcomes. *Eur J Surg Oncol*. 2024;50(4):108018. doi:10.1016/j.ejso.2024.108018
19. Volinsky-Fremond S, Horeweg N, Andani S, et al. Prediction of recurrence risk in endometrial cancer with multimodal deep learning. *Nat Med*. 2024;30(7):1962–1973. doi:10.1038/s41591-024-02993-w
20. Švecová M, Dubayová K, Birková A, Urdžik P, Mareková M. Non-invasive endometrial cancer screening through urinary fluorescent metabolome profile monitoring and machine learning algorithms. *Cancers*. 2024;16(18):3155. doi:10.3390/cancers16183155
21. Goyal M, Tafe LJ, Feng JX, et al. Deep learning for grading endometrial cancer. *Am J Pathol*. 2024;194(9):1701–1711. doi:10.1016/j.ajpath.2024.05.003
22. Njoku K, Pierce A, Chiasserini D, et al. Detection of endometrial cancer in cervico-vaginal fluid and blood plasma: leveraging proteomics and machine learning for biomarker discovery. *EBioMedicine*. 2024;102:105064. doi:10.1016/j.ebiom.2024.105064
23. Liu W, Ma J, Zhang J, et al. Identification and validation of serum metabolite biomarkers for endometrial cancer diagnosis. *EMBO Mol Med*. 2024;16(4):988–1003. doi:10.1038/s44321-024-00033-1
24. Wang CW, Muzakky H, Firdi NP, et al. Deep learning to assess microsatellite instability directly from histopathological whole slide images in endometrial cancer. *NPJ Digit Med*. 2024;7(1):143. doi:10.1038/s41746-024-01131-7
25. Kahaki S, Hagemann IS, Cha KH, et al. End-to-end deep learning method for predicting hormonal treatment response in women with atypical endometrial hyperplasia or endometrial cancer. *J Med Imaging*. 2024;11(1):17502. doi:10.1117/1.JMI.11.1.017502
26. Swarnkar B, Maheshwari P, Khare N, Gyanchandani M. “Early diagnosis of endometrial cancer: an ensemble-based deep learning approach,” in *2024 IEEE International Students’ Conference on Electrical, Electronics and Computer Science (SCEECS)*, IEEE, 2024, pp. 1–7.
27. Yasar S, Yagin FH, Melekoglu R, Ardigò LP. Integrating proteomics and explainable artificial intelligence: a comprehensive analysis of protein biomarkers for endometrial cancer diagnosis and prognosis. *Front Mol Biosci*. 2024;11:1389325. doi:10.3389/fmolb.2024.1389325
28. Liu J, Hu D, Lin Y, et al. Early detection of uterine corpus endometrial carcinoma utilizing plasma cfDNA fragmentomics. *BMC Med*. 2024;22(1):310. doi:10.1186/s12916-024-03531-8
29. Sawant A, Kulkarni S, Sawant M. Ultrasound super resolution imaging for accurate uterus tumor detection and malignancy prediction. *J Pharm Biomed Anal*. 2024;3:100029. doi:10.1016/j.jpba.2024.100029
30. Swarnkar B, Khare N, Gyanchandani M. enhancing endometrial tumor detection: early diagnosis with advanced vision transformer architecture. in *2024 IEEE International Students’ Conference on Electrical, Electronics and Computer Science (SCEECS)*, IEEE, 2024, pp. 1–6.
31. Umamoto M, Mariya T, Nambu Y, et al. Prediction of mismatch repair status in endometrial cancer from histological slide images using Various deep learning-based algorithms. *Cancers*. 2024;16(10):1810. doi:10.3390/cancers16101810
32. Liu X, Qin X, Luo Q, et al. A transvaginal ultrasound-based deep learning model for the noninvasive diagnosis of myometrial invasion in patients with endometrial cancer: comparison with radiologists. *Acad Radiol*. 2024;31(7):2818–2826. doi:10.1016/j.acra.2023.12.035
33. Singh TG, Karthik B, Wahengbam M. Utilizing A Deep Learning Model’s Hotspot Detection Technique for Uterine Cancer Image Segmentation,” in *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*. 2024;1–6.
34. Xue X, Liang D, Wang K, et al. A deep learning-based 3D Prompt-nnUnet model for automatic segmentation in brachytherapy of postoperative endometrial carcinoma. *J Appl Clin Med Phys*. 2024;25(7):e14371. doi:10.1002/acm2.14371
35. Feng L, Chen C, Wang L, et al. ECMS-NET: a multi-task model for early endometrial cancer MRI sequences classification and segmentation of key tumor structures. *Biomed Signal Process Control*. 2024;93:106223. doi:10.1016/j.bspc.2024.106223
36. Schonfeld SJ, Hartge P, Pfeiffer RM, et al. An aggregated analysis of hormonal factors and endometrial cancer risk by parity. *Cancer*. 2013;119(7):1393–1401. doi:10.1002/cncr.27909
37. Vicky M, et al. Endometrial cancer (primer). *Nat Rev Dis Primers*. 2021;7(1).
38. Epstein E, Fischerova D, Valentin L, et al. Ultrasound characteristics of endometrial cancer as defined by International Endometrial Tumor Analysis (IETA) consensus nomenclature: prospective multicenter study. *Ultrasound Obstet Gynecol*. 2018;51(6):818–828. doi:10.1002/uog.18909
39. Shorten C, Khoshgofaar TM. A survey on image data augmentation for deep learning. *J Big Data*. 2019;6(1):1–48. doi:10.1186/s40537-019-0197-0
40. ElGhany SA, Ibraheem MR, Alruwaili M, Elmogy M. Diagnosis of various skin cancer lesions based on fine-tuned resnet50 deep network. *Comput Mater Continua*. 2021;68(1).
41. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:3113–3121.
42. Panda MK, Sharma A, Bajpai V, Subudhi BN, Thangaraj V, Jakhetiya V. Encoder and decoder network with ResNet-50 and global average feature pooling for local change detection. *Comput. Vision Image Understanding*. 2022;222:103501.
43. Theckedath D, Sedamkar RR. Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks. *SN Comput Sci*. 2020;1(2):79. doi:10.1007/s42979-020-0114-9
44. Qiao S, Wang H, Liu C, Shen W, Yuille A. Micro-batch training with batch-channel normalization and weight standardization. *arXiv preprint arXiv:1903.10520*. 2019.
45. Yuan L, et al. “Tokens-to-token vit: training vision transformers from scratch on imagenet,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 558–567.
46. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M. Transformers in vision: a survey. *ACM Comput. Surv*. 2022;54(10s):1–41. doi:10.1145/3505244
47. Wu K, Peng H, Chen M, Fu J, Chao H. “Rethinking and improving relative position encoding for vision transformer,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 10033–10041.
48. Chen M, et al. Searching the search space of vision transformer. *Adv Neural Inf Process Syst*. 2021;34:8714–8726.
49. Li Y, Zhang K, Cao J, Timofte R, Van Gool L. Localvit: bringing locality to vision transformers. *arXiv preprint arXiv:2104.05707*. 2021.
50. Sharma AK, Nandal A, Dhaka A, et al. HOG transformation based feature extraction framework in modified Resnet50 model for brain tumor detection. *Biomed Signal Process Control*. 2023;84:104737. doi:10.1016/j.bspc.2023.104737
51. Zhang P, et al. “Multi-scale vision longformer: a new vision transformer for high-resolution image encoding,” in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 2998–3008.

52. Wang J, Yu X, Gao Y. Feature fusion vision transformer for fine-grained visual categorization. *arXiv preprint arXiv:2107.02341*. 2021.
53. Tan H, Liu X, Yin B, Li X. MHSA-Net: multihead self-attention network for occluded person re-identification. *IEEE Trans Neural Netw Learn Syst*. 2022;34(11):8210–8224. doi:10.1109/TNNLS.2022.3144163
54. Hu W, Hu M, Liu F, Han Y. “P-vit: a simplified vision transformer model based on FFN and simple attention,” in *International Conference on Knowledge Science, Engineering and Management*, Springer, 2024, pp. 316–326.
55. Wahid JA, Mingliang X, Ayoub M, Hussain S, Li L, Shi L. A hybrid resnet-vit approach to bridge the global and local features for myocardial infarction detection. *Sci Rep*. 2024;14(1):4359. doi:10.1038/s41598-024-54846-8
56. Hossain MB, Iqbal SMHS, Islam MM, Akhtar MN, Sarker IH. Transfer learning with fine-tuned deep CNN ResNet50 model for classifying COVID-19 from chest X-ray images. *Inform Med Unlocked*. 2022;30:100916. doi:10.1016/j.imu.2022.100916
57. Basha SHS, Dubey SR, Pulabaigari V, Mukherjee S. Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing*. 2020;378:112–119. doi:10.1016/j.neucom.2019.10.008
58. Liu W, Wen Y, Yu Z, Yang M. Large-margin softmax loss for convolutional neural networks. *arXiv preprint arXiv:1612.02295*. 2016.
59. Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. *J Big Data*. 2016;3:1–40. doi:10.1186/s40537-016-0043-6
60. Jiang Y, Wang C, Zhou S. Artificial intelligence-based risk stratification, accurate diagnosis and treatment prediction in gynecologic oncology. In: *Seminars in Cancer Biology*. Elsevier; 2023:82–99.
61. Gulia C, Signore F, Gaffi M, et al. Y RNA: an overview of their role as potential biomarkers and molecular targets in human cancers. *Cancers*. 2020;12(5):1238. doi:10.3390/cancers12051238
62. Yue W, Han R, Wang H, et al. Development and validation of clinical-radiomics deep learning model based on MRI for endometrial cancer molecular subtypes classification. *Insights Imaging*. 2025;16(1):107. doi:10.1186/s13244-025-01966-y
63. Qi X. Artificial intelligence-assisted magnetic resonance imaging technology in the differential diagnosis and prognosis prediction of endometrial cancer. *Sci Rep*. 2024;14(1):26878. doi:10.1038/s41598-024-78081-3
64. Coad CA, Santoro M, Zybin V, et al. A radiomic-based machine learning model predicts endometrial cancer recurrence using preoperative CT radiomic features: a pilot study. *Cancers*. 2023;15(18):4534. doi:10.3390/cancers15184534
65. Altal OF, Sindiani AM, Amin M, et al. Hybrid attention-enhanced mobilenetv2 with particle swarm optimization for endometrial cancer classification in ct images. *Inform Med Unlocked*. 2025;57:101662. doi:10.1016/j.imu.2025.101662

International Journal of Women's Health

Publish your work in this journal

The International Journal of Women's Health is an international, peer-reviewed open-access journal publishing original research, reports, editorials, reviews and commentaries on all aspects of women's healthcare including gynecology, obstetrics, and breast cancer. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-womens-health-journal>

Dovepress
Taylor & Francis Group