International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V74I2P109

Volume 74 Issue 2, 125-150, February 2026

Original Article

An Agentic AI Model for MRI-Based Staging and Grading of Endometrial Cancer using MedSAM and ViT

Sumitha B S[1], Mohammed Tajuddin[2], Vikram Patil[3], Shweta R Poojary[4], Aarathi Santhosh[5]

[1,2]*Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, Karnataka, India.*
(affiliated to Visvesvaraya Technological University (VTU), Belgaum, Karnataka, India.)
[1]*Department of Information Science and Engineering, Aditya College of Engineering and Technology, Bangalore, Karnataka, India. (affiliated to Visvesvaraya Technological University (VTU), Belgaum, Karnataka, India.)*
[3,4]*Department of Radiology, JSS Medical College, JSS Academy of Higher Education and Research, Mysuru, Karnataka, India.*
[5]*R I Consultant Obstetrician - Gynecologist I Narayana Hospital, Mysore, Karnataka, India.*

[1]*Corresponding Author : bssumitha@gmail.com*

Received: 28 July 2025 Revised: 10 January 2026 Accepted: 20 January 2026 Published: 14 February 2026

Abstract *- Endometrial Cancer (EC) is one of the most prevalent gynecological cancers globally. Accurate preoperative staging and the subsequent tumor grading are crucial for prognosis, surgical decisions, and treatment plans. Generally, for MRI evaluation, the radiologists depend totally on their own interpretation, which will be a mixture of subjectivity, a lengthy process, and inconsistency among different observers. The problems outlined above are addressed by creating a highly capable Artificial Intelligence (AI) framework that relies on multiple MRI modalities and clinical attributes for the precise identification of Deep Myometrial Invasion (DMI) and tumor grading. For accurate anatomical segmentation, the framework makes use of the Medical Segment Anything Model (MedSAM). After that, a Uterine Cavity Line Generation Algorithm (UCLGA) is employed to determine the depth of myometrial invasion. The Vision Transformer (ViT) model, fine-tuned via Low-Rank Adoption (LoRA), is used for feature representation and learning. Simultaneously, the agentic reasoning module successively enhances prediction through self-reflection and clinical knowledge. Furthermore, multiple smart agents are applied for segmentation visualization, report generation, compliance monitoring, and scheduling to provide a modular and interpretable system. The model is tested on the EC-MRI dataset, and its overall accuracy reached 96.82%, sensitivity 95.85%, precision 96.77%, F1 score 95.31%, and specificity 96.64%. The model surpassed the other current models in accuracy. Overall, the findings imply that this model is a clinically significant, elucidative, and efficient AI system that could support oncologists and radiologists in the preoperative EC evaluation.*

Keywords *- Agentic AI, Endometrial Cancer, LoRA, MRI, Myometrial Invasion, MedSAM, Tumor Grading, UCL, Vision Transformer.*

1. Introduction

Endometrial cancer, also known as uterine corpus cancer, is considered the most common malignancy of the female reproductive system in developed countries and has a worldwide increase in the occurrence rate. This malignancy predominantly affects postmenopausal and perimenopausal women. The contemporary lifestyle contributes to the escalating incidence of obesity and increases the chance of EC. The yearly risk of mortality for EC patients shows an increasing pattern [1]. Surgery is generally the most common treatment for EC, encompassing bilateral salpingo-oophorectomy, total hysterectomy, and lymph node evaluation [2]. Reports from 1990 to 2021 indicate a notable worldwide increase in EC cases among those aged 55 and above. During these 30 years, the cases doubled, rising to 360,253 in 2021 from 141,173 in 1990. The occurrence rate per 100,000 people increased to 45.81 from 39.22, underscoring an increasing health issue. In 2020, more than 417,000 women globally were diagnosed with endometrial cancer, indicating a 132% increase over the last three decades. EC instances are projected to increase by around 40% from 2020 to 2040. By 2036, the worldwide occurrence of EC in postmenopausal individuals aged 55 and the elderly is anticipated to increase by 6.5%, although the fatality rate is forecasted to decrease by 8%. Gynecologic malignancies are a significant worldwide health issue, representing almost 15% of newly identified tumor cases and fatalities amongst the female population in 2020. The cervical cancer constituted the most serious gynecological cancer at 7.70%, then the ovarian cancer at 4.7%, uterine cancer at 2.2%, and vulvar and vaginal malignancies [3]. GLOBOCAN estimations indicate that in 2022, there were approximately 1,473,427 new instances of gynecologic tumors and 680,372 fatalities globally [4]. The common diagnostic methods for EC, as depicted in Figure 1, are Magnetic Resonance Imaging (MRI), Ultrasound (US), histopathology, and hysteroscopy, which provide crucial information for clinical decision-making. The assessment of these methods depends largely on the skill of pathologists and radiologists. It is a process that was not only time-consuming and labor-intensive but also prone to inaccuracies due to human factors. As a result of this interdependence, diagnostic accuracy may be reduced, and the initiation of critical therapies may be delayed, both of which would have negative implications for the patient's health [5].

Early diagnosis is essential for the management and prognosis of EC. Conventional screening protocols additionally employ techniques such as Computed Tomography (CT) and Positron Emission Tomography (PET) [6]. DMI is significant in the preoperative assessment of EC. For DMI evaluation, a pelvic MRI or a transrectal or transvaginal US conducted by a specialist is suggested. Fertility-sparing treatments could be suggested for patients with grade one, stage IA endometrioid EC, if there are no MI and no additional risk factors are present. The evaluation of MRI or transvaginal US in ascertaining the absence of MI or shallow MI relies on extrapolation from data concerning the diagnosis of DMI [7].

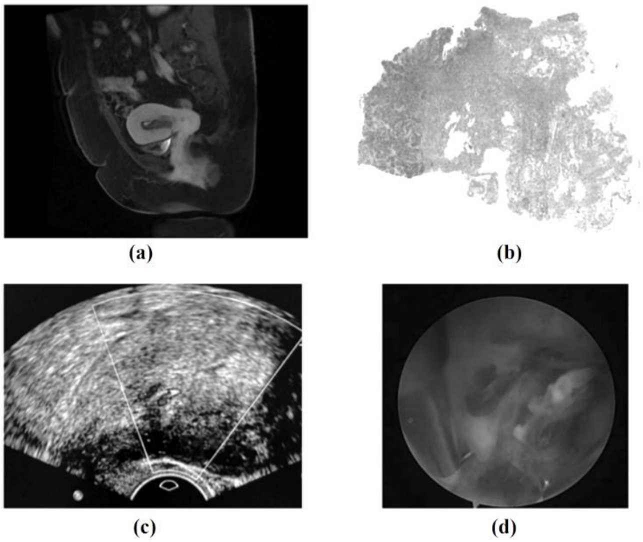

(a) (b) (c) (d)

Fig. 1 Imaging modalities for EC: (a) MRI, (b) Histopathology, (c) US, and (d) Hysteroscopy.

1.1. FIGO Staging for EC

The International Federation of Gynecology and Obstetrics (FIGO) staging method indicates that 75% of cases are diagnosed at stage I, with endometrial carcinoma further categorized into stages IA and IB based on the extent of MI. DMI, characterized by infiltration depth of ≥50% of the myometrial thickness, is regarded as a critical prognostic marker in endometrial carcinoma, as tumors exhibiting DMI possess an increased likelihood of pelvic lymph node and paracervical invasion metastases [8]. The FIGO staging system is employed to ascertain the surgico-pathological staging of endometrial carcinoma [9].

Table 1. Stages of EC based on FIGO-2023 [8]

FIGO Stages	Stage Classes with Description
Early Stage	
IA1	Type I, on the polyp or endometrium
IA2	Type I, MI<50%
IA3	Type I, uterus + 1 ovary (intact capsule)
IB	Type I, MI≥50%
IC	Type II, on polyp or endometrium (MI–)
IIA	Type I, cervical stroma
IIB	Type I, LVSI+
IIC	Type II, MI+
Advanced Stage	
IIIA1	Adnexa (direct or metastasis) - excluding IA3
IIIA2	Uterine serosa
IIIB1	Vagina or parametria (direct or metastasis)
IIIB2	Pelvic peritoneum implants
IIIC1	Pelvic lymph node metastasis
IIIC2	Infrarenal para-aortic lymph node metastasis
IVA	Bladder or intestinal mucosa
IVB	Abdominal peritoneum implants
IVC	Distant metastasis (lung, liver, bone...) or distant lymph node metastasis (inguinal...)

Note: Type I (non-aggressive histology): low-grade (G1/G2) endometrioid; Type II (aggressive histology): high-grade endometrioid (G3), serous, clear cell, carcinosarcoma, mesonephric, neuroendocrine, gastro-intestinal mucinous, mixed, undifferentiated; LVSI: lymphovascular space involvement. Molecular Subtyping Implications: I–II POLEmut → IAmPOLEmut (MI+/–, LVSI +/–, Type I or II); I–II p53abn– → IICmp53abn (MI+/–, LVSI +/–, Type I or II).

FIGO classifies Grade 1 tumors as well-differentiated, resembling normal tissue, and typically exhibiting a positive prognosis. Grade 2 cancers possess a dense feature that varies from 6% to 50% and are categorized as differentiated. Grade 1–2 endometrial carcinomas are categorized as type I; grade 3 tumors with a dense feature over 50% are characterized as high grade and poorly differentiated, exhibiting characteristics distinct from normal endometrial tissue, demonstrating aggressive behaviour, and correlating with a poor prognosis. Grade 3 Endometrial Carcinoma (EC) is categorized as type II EC, typically affects postmenopausal women, and is not linked to endocrine abnormalities [10]. The different stages of EC classified according to FIGO-2023 staging are presented in Table 1 [11]. MRI is essential for treatment planning, since it offers critical insights into tumor stages, including dimensions and invasion depth in the cervical stroma and myometrium, as well as involvement of pelvic lymph node status and anatomical structures. For staging EC, FIGO employs a surgical staging approach; however, expert consensus and recent studies suggest that MRI assessment prior to treatment is necessary to determine the most suitable therapy [12].

1.2. AI in Cancer Diagnosis

In nuclear medicine and radiology, the medical applications of AI systems show potential for several tasks, such as tumor detection, image reconstruction, and the development of diagnostic biomarkers. This serves as prognostic assessment, diagnostic assistance, and treatment efficacy prediction in gynecological oncology [13]. The integration of radiomic and Deep Learning (DL) models into pre-operative imaging (US, MRI) provides non-invasive alternatives for patient staging [14]. Recent advancements in AI present novel prospects to improve diagnostic accuracy and therapeutic decision-making. AI is increasingly recognized in gynecological oncology, where it is utilized to enhance both diagnostic and therapeutic strategies. Although endometrial carcinoma is the primary emphasis, recent advancements have also demonstrated potential in other gynecological malignancies. AI is explored as an implementation method for early, non-invasive identification of ovarian cancers and for the analysis of histopathology images in predicting the molecular types of EC malignancies. Expanding the range of AI applications may improve diagnosis accuracy and result in the detection of gynecological cancers [15].

1.3. Problem Statement

Several clinical challenges exist for medical imaging of EC for staging and grading preoperatively, despite advances in medical imaging and AI. Conventional diagnostic processes require the interpretation of MRI scans that are manually performed by radiologists and, thus, are subjective and time-consuming, possessing interobserver variability when evaluating DMI and tumor grading. Second, most of the deep learning techniques do not consider clinical data and never provide explainable output or simulate real-world diagnostic reasoning, which makes the decision-making process delayed and inaccurate, potentially leading to under- or overtreatment. Hence, the need arises for a robust, explainable, and clinically adaptable AI framework that autonomously analyzes multimodal MRI and clinical data, enhances the accuracy, reliability, and efficiency of EC staging and grading, and assists clinicians with unambiguous evidence-based recommendations.

1.4. Research Gap and Novelty

The significance of this research focuses on resolving the gaps concerning the accuracy and computation of preoperative staging and grading in EC, where treatment outcomes greatly depend on early and accurate diagnosis. Whereas in the general approaches, manual MRI interpretation dominates, this research proposes a novel Agentic AI framework that autonomously integrates medical images and clinical data into a single diagnostic system. Thus, its novelty lies in the multi-agent architecture that combines MedSAM for high-accuracy segmentation, the UCL generation algorithm for anatomical assessments, and a ViT fine-tuned with LoRA for feature extraction and classification. Finally, the system improves predictions through an agentic reasoning module that uses self-reflection and clinical knowledge and offers tumor board-style reports.

1.5. Research Objectives

The research hypothesis is that the proposed agentic AI model will not only be able to outperform others in diagnosis but will also be more interpretable, as a result of providing end-to-end automation, anatomical reasoning, and iterative clinical knowledge integration. Furthermore, it will lead to a decrease in variations among observers and subsequently upgrade the support of decision-making for radiologists and oncologists. The key objectives of this work are discussed as follows.

- To develop an Agentic AI-based framework that integrates MRI imaging and clinical data for the precise diagnosis and staging of EC.
- To implement MedSAM for high-precision segmentation of anatomical regions such as the uterus, myometrium, and tumor in MRI scans.
- To apply the UCLGA for measuring the depth of invasion within the myometrium and further FIGO staging.
- To use a ViT model fine-tuned using LoRA for proficient visual feature extraction and tumor grade classification.
- To integrate an agentic reasoning module supporting iterative self-reflection and external clinical knowledge absorption to refine diagnostic predictions.
- To develop assistant intelligent agents for report generation, compliance auditing, segmentation visualization presentations, and scheduling, assisting in clinical decision-making.
- To evaluate the performance based on standard metrics with accuracy, sensitivity, precision, F1-score, and specificity on a curated dataset of EC cases.
- To compare the proposed framework to state-of-the-art models, thereby proving its superiority over them both in terms of diagnostic accuracy and explainability.
- To conclude and discuss the advantages and limitations of the developed model with further improvements.

The paper is organized as follows: the first section includes a review of the recent research works proposed for EC detection and classification. The next section presents the detailed modelling and implementation of the proposed MedSAM-ViT model. Then, the subsequent section presents the experimentation analysis and results of the model and a comparison with current models. At last, the research concludes with the findings and recommendations for future research initiatives.

2. Literature Review

This section reviews the recent models developed for the staging and classification of EC and depth of MI using MRI with various state-of-the-art techniques. Table 2 highlights the critical analysis of the reviewed models with their advantages and limitations. A computer-aided diagnostic method utilizing a multi-stage DL model was proposed in [16] to diagnose early EC on sagittal T2-Weight (T2W) images by assessing the level of MI. The Single Shot multibox Detector (SSD)-based prediction method and an Attention-based U-Net segmentation method were developed to choose, crop, and segment MRI images. The ellipse fitting technique was employed to create a UCL for the determination of the depth of the MI for classification. In the independent testing dataset, the model for detecting the tumor and uterus attained average precision rates of 98.70% and 87.93%, respectively.

An automated segmentation framework based on DL using a U-net architecture was developed in [17] to delineate the tumor and uterus in MRI images. A semantic segmentation model based on the U-Net architecture was trained to delineate the tumor and uterine regions in MRI images. The Tumor-to-Uterine Region (TUR) area ratio was subsequently computed from the segmentation map. The stage IA or IB of EC patients was determined using TUR, and the outcomes of the patient's pathological diagnosis, which identified the ideal staging thresholds for stages IA and IB.

A prediction model in [18] that utilized diffusion-weighted imaging features derived from DL and radiomics, in conjunction with clinical data and Apparent Diffusion Coefficient (ADC) values, to detect microsatellite instability in EC. Traditional radiomics features and DL features based on Convolutional Neural Networks (CNN) were extracted from DWI. Logistic Regression (LR) and Random Forest (RF) were utilized as classification algorithms. DL features, ADC values, clinical factors, radiomic features, and their combination were utilized to develop DL, ADC, clinical, radiomic, and integrated models. The findings indicated that the integrated approach led to enhanced risk classifications.

The efficacy of DL algorithms integrated with MRI for the risk assessment and prediction of EC was analyzed in [19]. Utilizing the DL convolutional neural network architecture known as residual networks with 101 layers (ResNet-101), channel attention and spatial attention modules were used to enhance the model's performance. The model demonstrated enhanced efficacy in identifying high-risk endometrial carcinoma, with increased sensitivity and specificity and exceptional predictive accuracy. The research in [20] integrated conventional radiomics with DL techniques to preoperatively assess the risk classifications of uterine endometrioid cancer and to create personalized treatments based on the classification. A Densenet121 model was employed to ascertain the ROI cross-section for DL image features. A Multinomial Logistic Regression (MLR) was employed to develop a multisequence conventional radiomics signature based on the selected traditional radiomics attributes. The results showed that the radiomics nomogram had the best diagnostic accuracy, and the results were highly significant.

The paper [21] presented a fully automatic method for the accurate detection of deep MI on MRI. One of the main contributions of the study was the introduction of the new geometric feature called LS, which quantitatively classified the irregularities of the uterine tissues caused by endometrial cancer (EC). The authors applied a feature fusion strategy that merged weak and strong features, in which several Probabilistic Support Vector Machines (PSVMs) were connected to geometric feature LS and texture features to derive more discriminative information, and the resulting PSVM models were then jointly pooled to create the ensemble PSVM model EPSVM. The results pointed out that persistence to the EPSVM model was certainly superior in terms of failed cases and true positives, or in other words, specific and sensitive performance.

The research outlined in [22] explored the effectiveness of the ML radiomics model that utilized multiparametric MRIs in discriminating stage IA EC from the benign endometrial lesions, and further investigation was made into the possible use of different combinations of models, which included the clinical factors along with the radiomic features. After performing data dimensionality reduction along with feature selection, nine different ML algorithms were used which included Logistic Regression (LR), Random Forest (RF), Support Vector Machines (SVM), K-Nearest Neighbours (KNN), Stochastic Gradient Descent (SGD), Extremely randomized Trees (ET), Light Gradient Boosting Machines (LightGBM), Decision Trees (DT), and eXtreme GB (XGBoost). The LR algorithm model was recognized as the best radiomics model, achieving the highest accuracy.

In [23], the effectiveness of DL in identifying Carcinosarcomas (CSs) and differentiating them from ECs using different MRI sequences was affirmed and authenticated. The DL model convolutional neural network (CNN) was learned with CS and EC for all the sequences and then validated. The results demonstrated that the DL model CNN showed diagnostic performance that was fair and even better than that of professional radiologists in detecting apart EC from CS on MRIs.

The research in [24] evaluated the supplementary benefit of Diffusion-Weighted Imaging (DWI) in comparison to Dynamic Contrast-Enhanced MRI (DCE-MRI) and T2-weighted Imaging (T2WI) for the preoperative assessment of MI in EC. The study underscored the importance of MRI for the preliminary staging of EC. It showed that incorporating DCE or DWI-MRI with T2WI markedly enhanced the diagnostic efficacy of MRI in evaluating the extent of MI in

EC. The study did not measure interobserver variability among radiology experts in evaluating MRIs.

An artificial intelligence model was implemented in [25] for the EC MRI imaging, significantly enhancing the accuracy of feature extraction from images and patient classification. The ResNet architecture was employed to enhance the network's performance. The addition of depth enhanced the model's accuracy, the properties of the network were upgraded repeatedly, and the damage functions of the reduced networks were derived. The detection accuracy of MRIs was evaluated utilizing three architectures: an optimized network, Res-Net, and a shallow CNN. The findings indicated that the specificity and accuracy of enhanced scanning and T2W imaging in MRI were 95% and 88.75% respectively. A Swin transformer-based methodology for the early diagnosis of EC with multi-body positions fusion was proposed in [26]. Utilizing the Swin transformer framework and its developed SW-MSA (shift windows multiple self-coherences) component, MRI images across the three areas (transverse, coronal, and sagittal) were cropped, improved, and classed, while fusion tests in these planes are conducted concurrently. The findings indicated that the enhanced Swin transformer model attained superior performance across all criteria in the EC multi-site fusion tasks.

The research in [27] assessed the precision of multiparametric MRIs (mpMRIs) in identifying MI depth in EC Associated with adenomyosis (EC-A) in comparison to EC with no adenomyosis, and examined the influence of various adenomyosis subgroups on the MI depth in EC. Adenomyosis could decrease the diagnostic efficacy of interpreting EC-MI. Nonetheless, there was no notable disparity in the efficiency of diagnosing MI depth between the endometrial carcinoma groups with and without concurrent adenomyosis, nor among the various subtypes of adenomyosis. The predictive efficacy of radiomics models derived from MRI for the risk classification and grading of early-stage EC was analyzed in [28]. The 3D radiomic features were derived from segmented EC images obtained from MRI images, with features taken from all three modalities. Subsequently, employing five-fold cross-validation alongside a multilayer perceptron (MLP) algorithm, these features were refined utilizing Pearson’s correlation coefficient to construct a predictive model for risk classification and grading of EC. Nevertheless, the amalgamation of all three sequences yielded improved predicted accuracy. The radiomics model of MRI possessed the capability to precisely predict risk classification and the early stages of EC. Two new models for predicting postoperative pathology based on MI of preoperative and grading in low-risk EC subjects were developed in [29]. Two predictive models, New Prediction Models (NPM1 & 2), were introduced. Both models were constructed based on the primary variables, the MI depth. In NPM1, the MI depth functioned as the primary variable. The model employed iterative imputation methods to rectify the inconsistencies identified in MI diagnosis outcomes. The second model eliminated the erroneous depth of MI data and employed labelled smoothing to enhance precision. The findings validated that NPM2 was a superior method to predict the groups compared to NPM1. The study in [30] developed CNNs for identifying EC by utilizing multiple sequences and cross sections, aiming at validating optimal CNN imaging instances and comparing their detection accuracy with that of expert radiologists. CNNs exhibited superior diagnostic efficacy for identifying EC via MRI. Despite the absence of significant changes, the inclusion of additional image types enhanced the diagnostic efficacy for certain individual image sets. The DL model employing CNNs demonstrated much superior performance with the axial contrast-enhanced T1-W images and single set images of axial apparent diffusion coefficients of water map in comparison to professional radiologists.

In [31], the DL techniques based on DenseNet121, ResNet101, and ResNet50 were applied to characterize deep transfer learning in predicting EC. The outputs of the multi-sequence model were consolidated through three decision-level fusion methods, and the best model was identified. A clinical model was then constructed by combining univariate and multivariate logistic regression studies to extract independent clinical factors. The multiparametric MRI model exhibits robust efficacy in preoperatively forecasting aggressive EC. Through the application and validation of an MDLR-multimodal DL radiomics model using MRI, the research in [32] has significantly advanced the preoperative differentiation of MI in EC cases. By utilizing the Integrated Sparse Bayes Extreme Learning Machines, a DL-Signature (DLS) was generated. Furthermore, by combining clinical features with DLS, a Clinical Model was created that encompassed both clinical attributes and the MDLR model. With DLS ensuring the integration of features, the MDLR model increased the preoperative accuracy in differentiating between the non-existence and existence of MI.

Table 2. Critical analysis of current research models

Ref	Approach	Applications	Advantages	Drawbacks
[16]	SSD-based detection, Attention U-Net, ellipse fitting for UCL	MI level detection on sagittal T2W MRI	High tumor/uterus detection precision (98.7% / 87.93%).	Complex multi-stage pipeline; lacks generalizability evidence.
[17]	U-Net for tumor and uterus segmentation, TUR area ratio	EC staging (IA vs IB)	Effective staging using the TUR metric.	Limited to 2-stage classification; does not integrate clinical data.

[18]	DL and radiomics features + ADC values + LR & RF	Microsatellite instability detection in EC	Multi-modal feature fusion improved risk classification.	Requires complex feature engineering; computationally heavy.
[19]	ResNet-101 + channel & spatial attention	High-risk EC identification	Enhanced sensitivity, specificity, and predictive power.	Interpretability of deep features is limited.
[20]	Densenet121 for ROI + traditional radiomics + MLR	Risk classification for personalized treatment	Superior diagnostic accuracy via radiomics nomogram.	No details on real-time applicability or processing time.
[21]	EPSVM using geometric (LS) + texture features	DMI detection on MRI	Superior specificity and sensitivity; good feature fusion.	Traditional ML-based models lack end-to-end automation.
[22]	Radiomics with 9 ML classifiers + clinical data	Distinguishing EC from benign lesions	LR outperformed others; robust combination models.	Manual feature selection; model comparison only in a 2-class setting.
[23]	CNN trained on multiple MRI sequences	EC vs CS classification	Outperformed radiologists in diagnosis.	Binary classification only; lacks staging or grading functionality.
[24]	Multi-sequence MRI (DWI + DCE-MRI + T2WI)	MI extent evaluation in EC	Enhanced MI assessment with image fusion.	Interobserver variability was not evaluated.
[25]	ResNet + deep network optimization	EC classification using MRI	High specificity (95%) and accuracy (88.75%).	Limited methodological transparency; the shallow CNN comparison is limited.
[26]	Swin Transformer with SW-MSA across planes	Multi-position fusion for early EC diagnosis	Superior performance across sagittal/coronal/transverse planes.	Requires high computational resources; lacks clinical data integration.
[27]	mpMRI analysis for EC-A vs EC cases	Evaluating the DMI impact of adenomyosis	Detailed subgroup analysis of the adenomyosis effect.	No significant advantage in diagnostic accuracy was found.
[28]	3D radiomics + MLP + 5-fold CV	EC grading and risk classification	High prediction accuracy; sequence fusion improved results.	Feature selection is highly dependent on correlation filtering.
[29]	NPM1 & NPM2 using MI depth and label smoothing	Postoperative pathology prediction	NPM2 improved precision via label smoothing.	Focused on low-risk EC; not applicable to broader EC cohorts.
[30]	CNN on multi-sequence & cross-section MRIs	EC detection vs radiologist performance	CNNs outperformed experts in certain sequences.	Lacks comprehensive staging and grading models.
[31]	DenseNet121, ResNet101, ResNet50 and logistic regression	Preoperative prediction of aggressive EC using multiparametric MRI	Multi-sequence fusion; integration of clinical factors; robust predictive performance.	CNNs lack global context, no anatomical segmentation, static fusion, and limited explainability.
[32]	MDLR multimodal DL radiomics with DLS and clinical feature fusion	Preoperative differentiation of MI in EC	Effective multimodal fusion; improved MI classification.	Handcrafted radiomics; binary MI only; no end-to-end automation; limited interpretability.

2.1. Research Gap Analysis

The extensive and methodical literature review has been discussed to present a detailed background on the subject of MRI-based EC staging and grading, with a systematic examination of existing DL, radiomics, transformer-based, and traditional ML approaches concerning their application scope, performance, automation, interpretability, and clinical integration. The proposed study offers novelty through a unified agentic AI framework, unlike previous research, which was confined to specific tasks such as binary staging, risk classification, or feature-driven prediction, and did not involve full automation or explainability of the procedure. The proposed framework is a combination of MedSAM-based anatomical segmentation, a UCLGA algorithm for accurate

MI quantification, and a ViT model that has been fine-tuned using LoRA for tumor grading. Moreover, the reviewed works do not utilize agentic reasoning, nor do they present multi-agent clinical support (reporting, compliance, visualization), or iterative self-reflection in accordance with clinical guidelines, which are the key points that set apart the proposed framework. The review comparison has placed the proposed framework as a novel and comprehensive solution that is also clinically applicable, and which has advanced the state of the art in preoperative endometrial cancer assessment, by clearly indicating the drawbacks of current approaches.

3. Materials and Methods

The developed research model implements an advanced agentic AI framework for use in the preoperative assessment of EC for deep MI detection and tumor grading. As shown in Figure 2, the workflow begins with the acquisition of an EC-MRI dataset, which involves gathering multi-sequence MRI data of patients affected by EC. The involved steps in preprocessing the MRI images are intensity adjustment, resolution normalization, and alignment with reference coordinate frames, so that they are of the best possible quality for subsequent procedures. Each corresponding data point then passes through the MedSAM, which is a generic medical image segmentation method that renders segmentation masks from an image encoder and a prompt encoder. MedSAM autonomously marks anatomical structures such as the endometrium, myometrium, and tumor region using bounding box prompts and image embeddings. From the segmentation output, the UCLGA obtains the UCL, which spatially references the depth of MI. The DMI calculation relies on depth measurement for tumor penetration (a) as relative to the total myometrial thickness (b). Hence, cases wherein ≥50% invasion is present are classed under deep MI. The anatomical attributes from the segmentation maps and tabular clinical features (patient age, hormonal results, or pathology diagnoses) are combined into a feature set. The features are then fed into a ViT that has been fine-tuned with LoRA. This method allows one to fine-tune parameters with low overhead. The ViT infers rich contextual representations from flattened image patches through self-attention operations in several transformer encoder blocks. The visual-semantic representation is forwarded to an MLP head that performs an initial classification into tumor grades of G1, G2, or G3. This prediction is successively refined in the agentic reasoning module, which is the base method of this study.

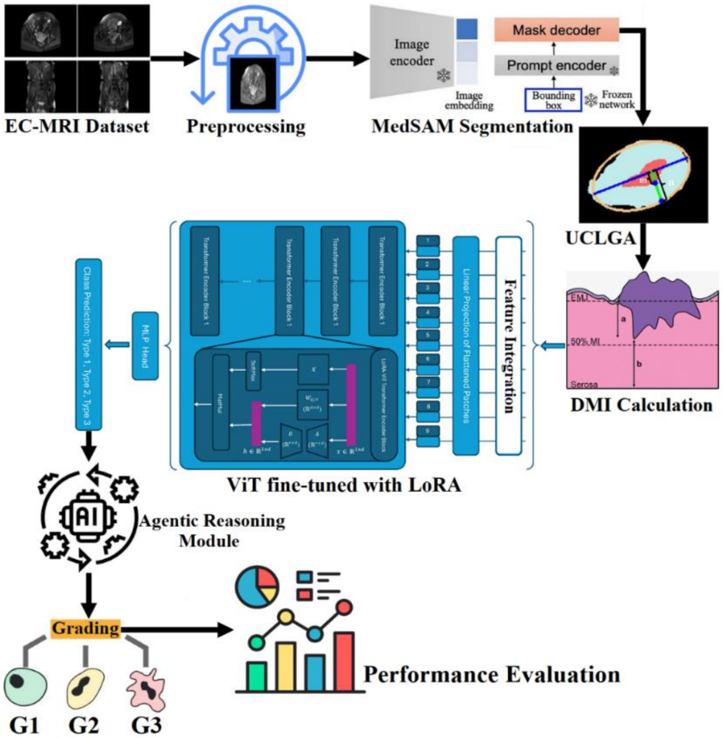


Fig. 2 Workflow of the developed research model

The standard performance metrics are used to evaluate the DMI presence and tumor grade prediction results even more. The depicted process shows how the combination of segmentation integration, anatomical computation, transformer-based learning, and agentic reasoning not only enhances the accuracy and reliability of EC staging and tumor stratification but also helps doctors make surgical and therapeutic decisions with greater confidence.

3.1. Dataset Details

An MRI dataset of EC was gathered from the Jagadguru Sri Shivarathreeshwara (JSS) Academy of Higher Education and Research and JSS Hospitals, Mysuru, Karnataka, India-570004, for this study. The dataset contains a total of 3500 MRI images showing different views-axial, coronal, and sagittal. Out of the total, 3100 images are found to be malignant, and 400 images are labeled as normal. The MRIs are obtained at a resolution of 256x256 pixels. The images selected from the acquired dataset are shown in Figure 3. Data summary of clinical information of EC patients from the collected dataset is shown in Table 3.

Table 3. Clinical information of the dataset

Parameter	Stage IA	Stage IB	p Value
No. of Patients	469	283	
Age	48.4 ± 8.9	61.8 ± 9.3	0.99
Endometrioid type			0.54
Grade 1	314	135	
Grade 2	141	116	
Grade 3	13	32	
Max. diameter (cm)			0.54
< 3	354	90	
≥3	115	193	
MI Depth			0.74
<50%	456	19	
≥50%	13	264	
Mixed Carcinoma			0.15
Yes	192	149	
No	277	134	

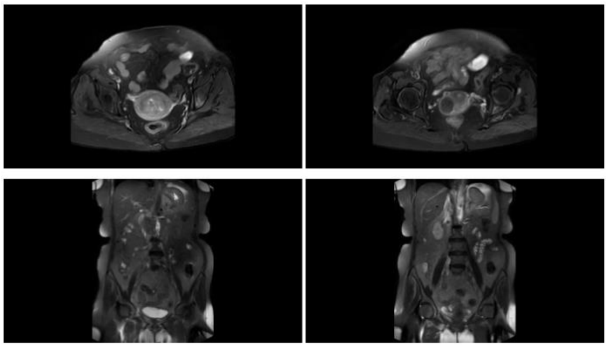


Fig. 3 Dataset Sample Images

3.1.1. MRI Image Description

Type: T2-weighted MRI (sagittal view). It is critical for MI assessment.

Annotations

- Tumor: Hypointense mass in the endometrium (outlined in red).
- Myometrial Invasion: Depth marked (<50% or ≥50%) with a yellow dashed line.
- Cervix: Stromal invasion indicated (green outline if present).
- Adnexa/Serosa: Abnormalities highlighted (blue outline).
- Lymph Nodes: Suspicious nodes (>1 cm) are circled in white.

Visualization: Image shows uterus with tumor extent, invasion depth, and annotations for FIGO staging (e.g., "Stage IB: ≥50% myometrial invasion").

3.2. Preprocessing

MRI images are often noisy, contain abnormalities, and have irrelevant information, which greatly reduces their quality. Therefore, pre-processing is necessary to eliminate the noise and artifacts that might affect the classification process while improving image quality and consistency. The current study uses down-sampling, filtering, enhancement, and augmentation as methods for MRI preprocessing.

3.2.1. Image Resizing

A digital image is a matrix with two dimensions M x N, where M represents the number of rows, while N represents the number of columns. Pixels are the finite number of discrete image elements that make up a digital image. The down-sampling method is used to speed up the data processing and make all the input images compatible with the ViT model at the same time. The original MRI dataset has a higher image resolution of 640x350, but it needs to be scaled down to the ViT model's input size of 224x224. The down-sampling process reduces the number of pixels in the image based on the sampling frequency rate. As a result, the image's dimensions and quality are reduced.

Interpolation is a technique through which the size of images is changed. The image interpolation has a bi-directional character, and its aim is to get the best possible estimate of a pixel's intensity and color based on the values of neighboring pixels. In the non-adaptive interpolation technique, an example being nearest neighbour (NN) interpolation, pixel manipulations are done directly without considering the features or the content of the image. Non-adaptive interpolation methods treat all pixels the same way, and therefore, they are easy to perform and require less processing power. The NN method is quite simple to comprehend, and it also consumes less time for processing than most other approaches. This method is also characterized as the point-shift method. It finds the pixel value that is closest to the surrounding coordinates of the specified interpolation point. In this way, the method will locate the closest corresponding pixel in the input image for every single pixel in the output image. With the increase in size, the pixels or color dots combine to form a new pixel. It generates edges that divide curves into angular segments or steps. This kind of interpolation impacts both the reduction and enlargement of images. An interpolation kernel for the NN technique is delineated as follows in equation (1).

$$f(x) = \begin{cases} 0, & x < 0.5 \\ 1, & x > 0.5 \end{cases} \tag{1}$$

The distance from the grid point to the interpolated point is represented by x [33].

3.2.2. Image Filtering

Digital images are particularly vulnerable to noise, which originates from inaccuracies in image acquisition and transmission in image processing. Images captured by MRI are generally vulnerable to salt-and-pepper noise, speckle noise, and Gaussian noise, which negatively impact image quality. The inferior quality of an image generally undermines the efficacy of subsequent activities, including feature extraction and reduction, and processed image classification. The noises must be eradicated before the further phases of processing. Consequently, a median filter is employed in this study for denoising purposes.

A median filter is a type of spatial filter that uses a moving window to replace the center value of the window with the median of the pixel values that are in the window. Filter removed noise but also obliterated fine details.

Researches mostly depend on median filters for its results, due to the filters' ability to provide quite good noise reduction, yet this is done without excessive blurring or similar effects. Basically, the filters are effective in different noise types. Median filters are widely used as smoothing techniques in signal/image processing and in time series analysis. A significant benefit of this filtering technique compared to linear filters is its ability to mitigate the influence of the input level of noise with an exceptionally high range. The median filtering is superior to the mean filtering as a non-linear filtering approach that effectively eliminates noise. It possesses the capability to eliminate 'impulse' noise (anomalous values that are either too low or too high). It is commonly asserted to be 'edge-preserving' as it supposedly maintains step edges without introducing blurring. Nonetheless, in the case of noise, it can marginally obscure edges in images.

$$Z_i = med\{WN_i\} = med\{Y_i + r : r \in WN\} \tag{2}$$

The fundamental median filter is defined as (2), whereas Y_i and Z_i represent the input and output at location i of the filter. The $[WN_i]_r r = 1, \dots, 2N+1$, represents the rth order statistic of the samples within the window. WN_i is $[WN_i]_1 < [WN_i]_2 < \cdots [WN_i]_{2N+1}$. In this research, the filter window size is 3x3 [34].

3.2.3. Image Enhancement

Incorporating the Quadrant Dynamic Histograms Equalization (QDHE) approach in this work could significantly enhance the identification and categorization of EC utilizing DL. QDHE improves image quality by addressing contrast changes and highlighting small details in MRI images, rendering key characteristics, such as lesions and abnormalities, more detectable. This improvement facilitates precise diagnosis and assists the deep learning model in recognizing complex patterns that could be challenging to identify. The QDHE process commences histogram segmentation utilizing the intensity values derived from the input histograms of the image. The histograms of the original image are first partitioned into dual sub-histograms. The mean values of the sub-histograms will serve as partitioning points to segment all the dual sub-histograms into 4 sub-histograms. The lower and upper extreme intensity values of the input histogram were employed as the initial and terminal points for the segments. There are similarities between the partitioning technique and the recursive sub-images HE utilized with the QDHE. The partitioning approach based on the median seeks to allocate pixel counts among sub-histograms evenly. The subsequent equation (3) could be employed to ascertain the position of each dividing point:

$$\left.\begin{aligned} b_1 &= 0.25 \times (A_H \times A_W) \\ b_2 &= 0.50 \times (A_H \times A_W) \\ b_3 &= 0.75 \times (A_H \times A_W) \end{aligned}\right\} \tag{3}$$

The intensities b_1, b_2 and b_3 are assigned indices of 0.25, 0.50, and 0.75, respectively. These values represent specified fractions of the overall pixel count in the input histograms. The variables A_H and A_W denote the height and width of the input image. The clipping process controlled the HE enhancement workflow, which resulted in the images that were not heavily brightened and that were even. The idea was to diminish the clipping effect by changing the input histogram symmetry. This was done by altering the histogram bin values according to the threshold value, which was related to the average image intensity values. The grey level dynamic range of each sub-histogram is calculated as the ratio between the total gray level spans and the total number of pixels in the respective sub-histograms. This method ensures that all the sub-histograms have the same amount of improvement space available. The whole procedure was expressed mathematically by the equation (4):

$$span_i = b_{i+1} - b_i \tag{4}$$

$$factor_i = span_i \times (\log_{10} B_i)^{\gamma} \tag{5}$$

$$range_i = (M-1) \times \frac{factor_i}{\sum_{b=1}^{4} factor_l} \tag{6}$$

In equation (4), $span_i$ denotes the dynamic grey level utilized by the ith sub-histograms of the input. In (5), b_i denotes the ith split point, while B_i signifies the total pixels in the sub-histogram. For the sub-histograms, the dynamic range in the resultant images is delineated as $range_i$, and the effect of B_i in (6) was measured with γ. The γ must be adjusted to precisely determine the level of all the sub-histograms in the output images. Nonetheless, the QDHE approach utilizes nearly equivalent overall pixel counts for every sub-histogram, rendering equation (5) inconsequential to the resultant dynamic ranges. To streamline QDHE and eliminate the term γ, the subsequent equation (7) could be restructured:

$$range_i = (M-1) \times \frac{span_i}{\sum_{b=1}^{4} span_l} \quad (7)$$

The dynamic level updated for the *i*th sub-histograms, as delineated by equations (8) and (9), is contained within the interval $[i_{start}, i_{end}]$.

$$i_{start} = (i-1)_{end} + 1 \quad (8)$$

$$i_{end} = i_{start} + range_i \quad (9)$$

The first value of i_{start} represents the least intensity in the updated assigned dynamic level. Upon determining the new dynamic levels for every quadrant sub-histogram, the concluding part of QDHE entails the independent equalization of all the sub-histograms. Once the histogram assigns the grey level $[i_{start}, i_{end}]$, the equalized outcome $h(x)$ for this segment, the transfer function described in equation (10) could be utilized.

$$t(x) = (i_{start} - i_{end}) \times d(X_l) + i_{start} \quad (10)$$

Thus, cumulative density functionality for the sub-histogram was denoted as $d(X_l)$ [35].

3.2.4. Data Augmentation

The data augmentation approach encompasses a broader range of solutions that function at the data level rather than at the architecture level of the model. It can increase the performance of DL models by synthetically producing diverse samples with balanced categories for the training dataset. A deep learning model demonstrates enhanced performance and accuracy when the dataset is sufficient in both quantity and quality. The training data must satisfy two criteria: adequate diversity and scale, both of which can be achieved by data augmentation. Important geometric operations, including random rotation, cropping, and flipping, are persistently in demand for data augmentation. They frequently increase the data volume to improve data diversity and are simple to implement. Geometrical data augmentation methods have demonstrated significant efficacy in increasing diversity and expanding data volume. Rotation is one of the basic geometric data augmentation methods. A set of images is created by rotating the original image at a specific angle, and the new images are then used as training samples in combination with the original images.

The EC MRI dataset consists of 3,500 images and is not sufficient for the model to be trained and tested properly. Therefore, the dataset is augmented using the geometric rotation data augmentation technique. The images are rotated by 45°, 90°, 135°, 180°, 225°, 270°, and 315°, as shown in Figure 4. As a result, the dataset is increased to a total of 28,000 images after geometric rotation, including 3,200 images of normal cases and 24,800 images of abnormal cases. The dataset was split in a ratio of 75:25.

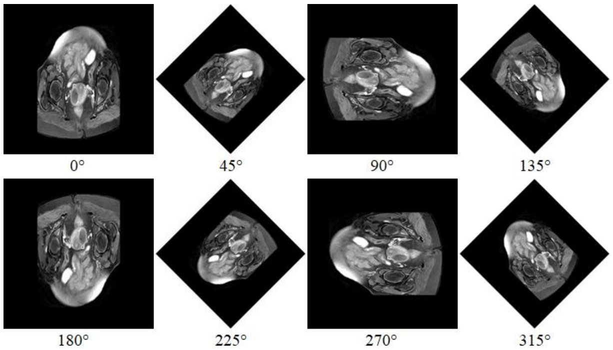


Fig. 4 Geometric Rotations Applied to the Images

3.3. MedSAM Segmentation

MedSAM is a model based on DL that was created for the purpose of segmentation of different anatomical lesions and structures within a variety of medical imaging sources. MedSAM was trained using a dataset of EC MRI scans. Its adjustable prompt mechanism gets the best possible balance between personalized and automatic control. Thus, making MedSAM a flexible instrument for the usual medical image segmentation. MedSAM is considered to be a basic model for full medical image segmentation. A key factor in creating such a model is the capacity to accommodate different imaging situations, anatomical characteristics, and clinical conditions. MedSAM is a segmentation method that allows the user to determine the segmentation targets by means of bounding boxes. In addition, drawing a bounding box is very useful, especially in situations where multi-object segmentation is needed. The study follows the network architecture proposed in SAM, which consists of an image encoder, a mask decoder, and a prompt encoder, as shown in Figure 5.

The image encoder maps the images given as input to a very large-dimensional embedded space. The prompt encoder turns the user-drawn box boundaries into feature representations with the help of positional encoding. After that, the mask decoder merges image embedding and swift features through cross-attention. The research network was built using a transformer model that has shown remarkable efficiency across a wide range of applications, including image recognition and natural language processing. The network deployed a ViT-based image encoding for feature extraction, a prompt encoder for the integration of user input via bounding boxes, and a mask decoder that provided confidence scores and segmentation results by using the prompt embedding, image embedding, and output tokens [36].

In order to attain a balance between segmentation and computational efficiencies, the study adopted the foundational ViT model as the picture encoder, since extensive assessments showed that bigger ViT models, such as ViT Large and ViT Huge, gave only slight improvements in accuracy while greatly increasing the demand on computational power. The basic ViT approach has 12 transformer layers, each made up of the Multilayer Perceptron (MLP) blocks and a multi-head self-attention block with layer normalization. Pre-training utilized masked auto-encoder modelling, succeeded by comprehensive supervised learning on the dataset. The input image (1024×1024×3) was transformed into a series of flattened two-dimensional patches measuring 16×16×3, resulting in a feature size of 64×64 in the image embedding after traversing the image encoder, which is downscaled by a factor of 16x.

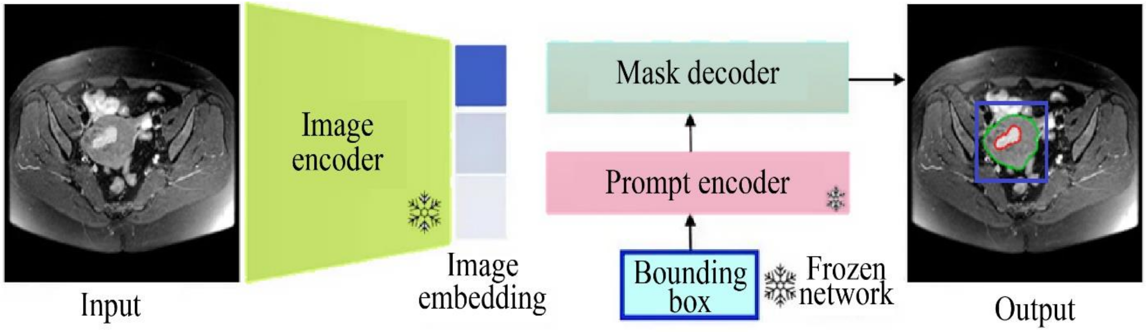


Fig. 5 Architecture of the MedSAM model

The prompt encoders translated the bounding box prompt's corner point into 256-dimensional vectors of embedded data. All the bounding boxes were denoted by the embedding pairs consisting of the bottom-right corner and top-left corner points. A simple mask decoder methodology was utilized to enable real-time interactions after computing the image embedding. It comprises two transformer layers for integrating the prompt encoding and image embedding, along with two inverted convolutional layers to elevate the embedded resolution to 256×256. Following that, the embedding was subjected to sigmoid activation and then bilinear interpolation following the input dimensions.

The work employed the unweighted sum of the cross-entropy loss and Dice loss as the last loss function, as it has demonstrated robustness across several medical image segmentation challenges. Let SM and GT represent the segmentation result and the ground truth, respectively. Let sm_i represent the voxel i's predicted segmentation and gt_i symbolize the ground truth, accordingly. N represents the voxel count in the image I. The binary cross-entropy loss was articulated as follows in Equation (11).

$$Ls_{BC} = -\frac{1}{N}\sum_{i=1}^{N}[gt_i \log sm_i + (1 - gt_i)\log(1 - sm_i)] \quad (11)$$

$$Ls_D = 1 - \frac{2\sum_{i=1}^{N} gt_i sm_i}{\sum_{i=1}^{N}(gt_i)^2 + \sum_{i=1}^{N}(sm_i)^2} \quad (12)$$

$$Ls = Ls_{BC} + Ls_D \quad (13)$$

The dice loss and final loss were articulated in the given Equations (12) and (13). To optimize the training process, the research presented the loss function as the unweighted aggregate of cross-entropy loss and Dice loss, thereby enhancing efficiency. MedSAM's exceptional generalization capability renders it a flexible solution for diverse medical image segmentation applications with enhanced efficiency and accuracy. The utilization of a unique technique to ascertain the depth of MI is a significant issue owing to the variability of uterine morphologies and cancer regions. Consequently, an approach for the automatic construction of UCL on a semantic segmentation map was developed to compute the MI depth. A line is derived as the simulated UCL. Subsequently, two maximal parallel lines to the UCL are established. One represents the maximal density of the myometrium to the UCL, while the other indicates the maximal extent of the cancer to the UCL. The proportion of the line's length is equivalent to the MI depth. The standard equation of an ellipse was shown in (14).

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \quad (14)$$

The research employed a direct least-squares method to fit the ellipse to independent data by reducing the algebraic distances, constrained by the equation $4ac - b^2 = 1$. It is simple to execute and highly resilient. Here, a, b, c, d, e, f represent the fitted ellipse variables derived from the pair of points (x, y) received from the inputs. The procedure was implemented on the uterine contour within the segmentation images, designating the long axis of the fitted ellipse as the UCL. Vertical lines are drawn at all the points of the UCL, and the distance ratio of every line perpendicular to the intersections of the cancer's borders and the uterine border was computed, with the maximal distance ratio designated as the depth of MI [37].

3.4. ViT-LoRA Model

ViT has achieved significant success and has proliferated across numerous vision applications, including semantic segmentation, image classification, and object identification. Pretrained ViT algorithms have been extensively utilized in downstream tasks, yielding exceptional outcomes via transfer learning. ViT is a novel DL method that takes the Transformer architecture, which was first used for Natural Language

Processing (NLP), and adapts it to the field of computer vision, especially to the classification of images. ViT is a novel method that treats images as patch sequences, just as word tokenization is done in language processing. That means the whole processing of the image is changed in a way that it can be considered as an advantage in the vision tasks, due to the Transformer's power in catching subtle relations among the sequences. Even though ViT models show high precision and better generalization over different tasks, their application in EC classification is still very difficult, particularly in clinical settings, because of the difficulties involved. This is mainly because ViT models, utilizing transformer architecture, include a substantially greater parameter count than earlier CNN-based models. Figure 6 depicts the architectural details of the implemented ViT-LoRA model [38].

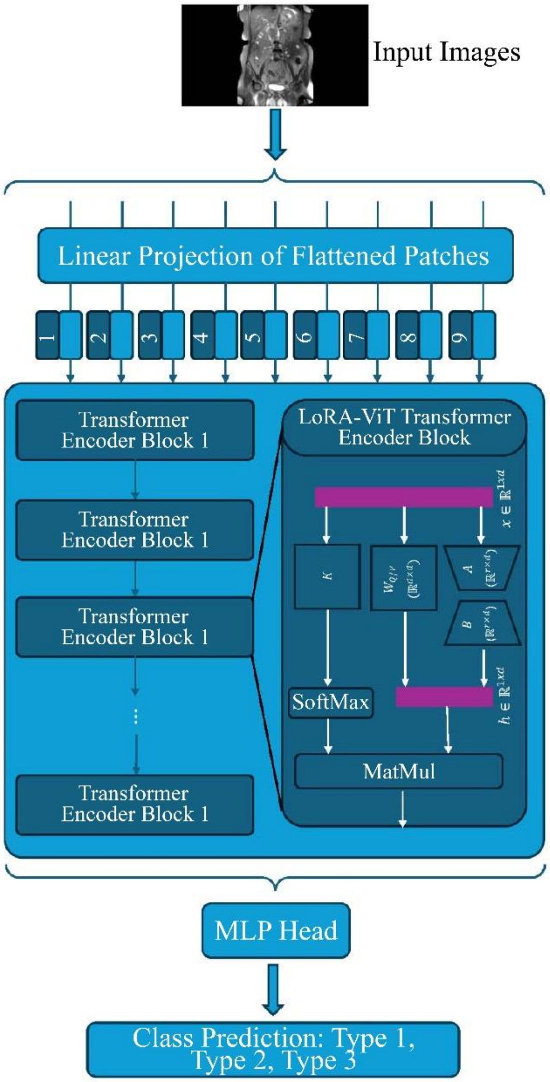


Fig. 6 Architecture of ViT-LoRA Model

Transformers have attained significant efficiency in NLP tasks, mostly due to their attention processes. The ViT has developed into a formidable architecture for the classification of images, building upon this notion. The ViT architecture consists of three essential components.

The image as input $x \in R^{H \times W \times C}$ is split into patches with fixed sizes to provide a sequential illustration of flattened 2D patches $x_p \in R^{N \times (p^2 \cdot C)}$, where H indicates the height of the image, W symbolizes the width of the image, C symbolizes the count of channels, and (P, P) specifies the resolution of all the image patches. The quantity of patches N could be determined as follows in equation (15).

$$N = \frac{H \times W}{P^2} \tag{15}$$

Prior to inputting the patch series into the Transformer, the linear projection (LP) was executed on the patches. In this LP, the patches of data were transformed into a D-dimensional vector space through multiplication with the embedding matrix E. The result of this LP was termed patch embedding. Positional embedding E_{pos} was added to the embeddings of the patch to allow the model to collect positional data in the images. Furthermore, the embedded image patches were combined with the class tokens learnable. x_{class}, which was crucial for the classification procedure. The preliminary patch embedding z_0, which includes the embedding image patches sequence and the class tokens, were calculated using equation (16):

$$z_0 = \left[x_{class}; x_p^1 E; E; \ldots; x_p^N E\right] + E_{pos},$$

$$E \in R^{N \times (p^2 \cdot C) \times D}, E_{pos} \in R^{(N+1) \times D} \tag{16}$$

Here, x_p^N denotes the image patch of n, $n \in 1,2, \ldots, N$.

The resultant embedded image patches were subsequently transmitted to the encoder of the Transformer. This encoder consists of L equivalent encoder blocks, each comprising two layers: a Multi-head Self-Attention (MSA) and a Fully Connected (FC) feed-forward Multi-Layer Perceptron (MLP) layer. In all the encoder blocks, the l-th layer obtains the input sequences from the preceding layers. z_{l-1}. The input z_{l-1} was subjected to normalization of layer, which standardizes the inputs with the dimension of features, enhancing training efficiency and performance. The result of the normalization layer was thereafter transmitted to the layer of MSA.

The result of the MSA was subsequently subjected to layer normalization once again. The result obtained from the layer normalizing was subsequently input into the layer of the MLP. Residual or skip connections were utilized in the encoding blocks to enhance the transmission of data within non-consecutive layers. These connections facilitate the propagation of gradients throughout the networks, unaffected by the non-linear activation function, hence mitigating the vanishing gradients problem. The gradients flow in the l-th encoding layer were delineated as follows in equations (17) and (18).

$$z_l' = MSA\big(LN(z_{l-1})\big) + z_{l-1}, \;\; l = 1, \ldots, L \qquad (17)$$

$$z_l = MLP\big(LN(z_l')\big) + z_l', \;\; l = 1, \ldots, L \qquad (18)$$

Here, LN indicates the normalization of the layer.

The MSA comprises layers such as linear, concatenation, self-attention, and concluding linear layers. In the MSA, numerous self-attention processes are executed simultaneously according to the count of heads k. In all heads, the D-dimension patch embeddings z were multiplied by the matrices of weights like U_q, U_k and U_v to derive the key (k), value (v), and query (q) matrices. The function of multiplication in all the heads was specified as in (19).

$$[q, k, v] = \left[zU_q, zU_k, zU_v\right], \quad U_q, U_k, U_v \in R^{D \times D_h} \qquad (19)$$

The derived matrices $q, k,$ and v are subsequently estimated into k subspaces, and the weighted average of every value V was computed. Attention weights were calculated in every head by evaluating the relation within every set of elements. (i, j), through the dot product of q^i and k^j. The resultant dot products signify the implication of patches in the sequences. The dot products of q and k were calculated using equation (20), followed by the application of a softmax layer to derive the weights for the values.

$$A = softmax\left(\frac{qk^T}{\sqrt{D_h}}\right), \quad A \in R^{N \times N} \qquad (20)$$

Here, $D_h = D/k$.

The matrix structures of the self-attention layers are concatenated and subsequently processed by a single linear layer with a trainable weight matrix. U_{msa}, yielding the following equation (21).

$$MSA(z) = [SA_1(z); SA_2(z); \ldots, SA_k(z)]U_{msa}, \quad U_{msa} \in R^{k.D_h \times D} \qquad (21)$$

The MSA's every head acquires information from numerous dimensions and perspectives, enabling the framework to encode complex properties simultaneously. The ViT model's classification is executed using an MLP, comprising two FC layers utilizing the activation function called Gaussian error Linear Units (GeLU). The GeLU assigns a weight to the input depending on its magnitude instead of its sign. In contrast to the ReLU, GeLU could yield both negative and positive outputs and demonstrates a greater range of curvature. This factor enabled the GeLU functions to more effectively approximate intricate functions in comparison to the ReLU. The visual depiction of the sequence is acquired using the following equation (22):

$$r = LN(z_L^0) \qquad (22)$$

In the encoding process, the final layer chooses the initial token of a sequence, z_L^0, and produces the representations of image r through the application of normalization of the layer. The resultant r was then processed by a compact MLP head, comprising a hidden layer utilizing the sigmoid functions for classification objectives.

LoRA, developed by Microsoft, offers an innovative solution to this challenge by modifying pre-trained vision methodologies for effective EC detection systems without necessitating full fine-tuning. This approach entails fixing the pre-trained model's weights and incorporating adaptable rank decomposition matrix structures into all the layers of the Transformer architecture. LoRA significantly decreases the number of parameters trainable in the process of fine-tuning. This process enhances the feasibility of the training function with standard medical imaging data sets while optimizing GPU resource utilization and storage, hence meeting essential clinical demands for space and speed. The work integrates weights of LoRA into all self-attention layers of a pre-trained ViT for the development of the model architecture. During fine-tuning, the changes to the pre-trained query matrix WQ and the value projection matrix W_V in a self-attention layer, the constraints are imposed through the incorporated weights of LoRA. The weights are represented using the low-rank decomposition and were articulated as follows in equation (23):

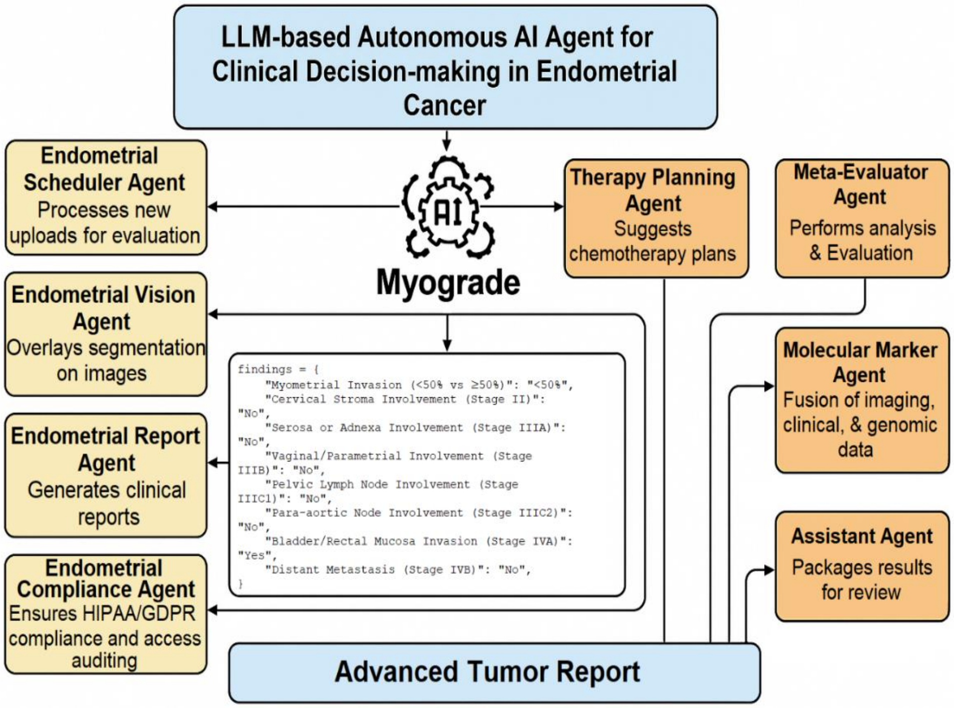
$$h = W_0 x + \Delta W_x = W_0 x + BAx \qquad (23)$$

Here, $x \in R^{H \times W \times C}$ represents the input, and $h \in R^{H \times W \times C}$ denotes the features of the output. The weight change ΔW is composed of two low-rank matrices, $B \in R^{d \times r}$ and $A \in R^{r \times d}$.

At the commencement of training, the work utilized an initialization of a random Gaussian for matrix A and initialized matrix B to zero. Thus, the product of matrices B and A, referred to as ΔW, is originally zero. The rank r of these low-rank matrices is significantly less than the model dimensions d, and the research experimentally fixed $r = 8$. Typically, r should not exceed 8, as the low-rank matrix amplification capacity is compromised when the rank reaches 64 in the experiment [39].

3.5. Agentic AI Reasoning Module

The Agentic AI reasoning module in this research represents a new multi-agent LLM-powered framework for clinical decision-making in endometrial cancer using multimodal MRI data. A special AI agent named "Myograde" lies at the core of this system, integrating LLM capabilities and domain-specific reasoning in staging, grading, and personalized reporting for endometrial carcinoma. The agentic system is built as a modular, extensible architecture designed for high-assurance medical use along multiple functional domains. The framework then initiates an Image Agent tasked with analyzing the DICOM/MRI, using segmentation outputs (MedSAM) to localize the tumor, uterus, and important anatomical landmarks, such as the UCL. Subsequently, from these, the system computes the depth of myometrial invasion, involvement of the cervical stroma, and other spatial features important for staging.



```
findings = {
    "Myometrial Invasion (<50% vs ≥50%)": "<50%",
    "Cervical Stroma Involvement (Stage II)":
"No",
    "Serosa or Adnexa Involvement (Stage IIIA)":
"No",
    "Vaginal/Parametrial Involvement (Stage
IIIB)": "No",
    "Pelvic Lymph Node Involvement (Stage
IIIC1)": "No",
    "Para-aortic Node Involvement (Stage IIIC2)":
"No",
    "Bladder/Rectal Mucosa Invasion (Stage IVA)":
"Yes",
    "Distant Metastasis (Stage IVB)": "No",
}
```

Fig. 7 Agentic AI-based reasoning module

As shown in Figure 7, the reasoning module continues with a network of collaborating sub-agents: the Endometrial Vision Agent superimposes the segmentation results on top of the images for visual verification; the Endometrial Report Agent generates structured clinical reports in patient's terms; the Endometrial Compliance Agent makes sure that all the agents in the system comply with HIPAA/GDPR, conduct access auditing, and the Endometrial Scheduler Agent constantly evaluates any new upload for processing.

In addition to imaging, lab indices (SII, PLR, NLR, platelet counts, etc.) and risk scores are integrated into the system, allowing for more refined staging classifications based on both radiologic and systemic biomarkers; it is equipped with fact-checking and clinical plausibility modules that simulate a tumor board review, contradicting what would be clinical integrity in profile. A Molecular Marker Agent predicts genomic markers such as ER, PR, MSI, POLE, and p53 based on a fusion of imaging, clinical, and genomic information. Consequently, on the basis of staging and molecular profiles, the Therapy Planning Agent offers personalized chemotherapy cycles, whereas, optionally, the Nutrition/Treatment Navigator takes care of monitoring for side effects, fatigue, and diet, thus supporting the patient through holistic care. The Meta-Evaluator Agent will compile insights across agents for the explainable AI outputs, including risk stratification and treatment recommendations, while maintaining verification guardrails and actionable safety warnings. The result would be an autonomous tumor board assistant with a human-in-the-loop interface, which, apart from prediction of DMI and tumor grades(G1-G3), can also generate explainable reports grounded in clinical practice and suggestions for next-step interventions. This agentic reasoning module sets the stage for a new paradigm in cancer diagnostics AI, going from static predictions to dynamic, context-aware clinical reasoning. In high-stakes clinical settings, Agentic AI facilitates distributed clinical reasoning by allocating responsibilities such as diagnosis, treatment planning, and vital monitoring to specialized agents. For instance, one agent may obtain patient history, another verifies findings against diagnostic criteria, and a third suggests therapy alternatives. These agents coordinate via shared memory and reasoning chains, providing coherent and secure suggestions. Applications encompass radiology assessment, ICU management, and pandemic responses. Despite the current lack of real-world applications owing to the new stage of the field, research indicates that Agentic AI has the potential to transform the healthcare sector [40].

3.6. Proposed MedSAM-ViT-LoRA Algorithm & Hyperparamters

The pseudocode for this developed research model is presented in the following. The hyperparameters of the developed model are presented in Table 4. The pseudocode for the proposed research begins with loading MRI scans and clinical data from the patient dataset.

Each image is preprocessed by resizing it to 224×224 pixels via NN interpolation, denoising through a median filter, enhancing with QDHE, and then geometrically augmenting the data using image rotations at four 45° increments to maximize the size of the dataset. Preprocessed images are then fed into the MedSAM model for segmentation of anatomical regions, such as the tumor, uterus, and myometrium.

Using the segmented maps, the UCL is generated through the UCLGA, with depths of MI calculated as tumor depth divided by total myometrial thickness and subsequently categorized as either shallow or deep invasion (≥50%). Simultaneously, image-derived features and clinical features are combined to form one unified feature set. These features are passed to the fine-tuned ViT model via LoRA to tokenize the inputs and compute initial tumor grade predictions (G1, G2, G3) through transformer layers and an MLP head.

The agentic reasoning module then improves upon these predictions by self-reflecting and consulting external clinical knowledge; the revised outputs should now be more accurate. The Endometrial Report Agent, Vision Agent, Scheduler Agent, and Compliance Agent then work together to create explainable reports, visualize the segmentation overlays, manage incoming data, and monitor compliance. The final outputs include DMI classification, tumor grade, and the complete AI-generated tumor board report.

Algorithm: MedSAM-ViT-LoRA Model

```
Initialization
Step 1: Load Patient Data
LOAD MRI_images, clinical_features
Step 2: Preprocessing
FOR each MRI_scan IN MRI_scans:
Resize image using Nearest Neighbour Interpolation
resized_image←nearest_neighbor_interpolation(MRI_scan, target_size=(224, 224))
Apply Median Filtering to remove noise
    denoised_image ← apply_median_filter(resized_image, window_size=3x3)
Enhance the image using QDHE

enhanced_image←QDHE_enhancement(denoised_image)
Data Augmentation via Rotation
    FOR angle IN [45, 90, 135, 180, 225, 270, 315]:
        rotated_image ← rotate_image(enhanced_image, angle)
        augmented_dataset.append(rotated_image)
    augmented_dataset.append(enhanced_image)
Step 3: Image Segmentation using MedSAM
FOR each preprocessed_scan:
    segmentation_map ← MedSAM(preprocessed_scan)
    STORE segmentation_map
Step 4: Generate Uterine Cavity Line using UCLGA
FOR each segmentation_map:
    UCL ← UCLGA(segmentation_map)
    STORE UCL
Step 5: Compute DMI (Deep Myometrial Invasion)
FOR each UCL, segmentation_map:
    a ← tumor_depth(segmentation_map, UCL)
    b ← total_myometrial_thickness(segmentation_map, UCL)
    DMI_ratio ← a / b
    IF DMI_ratio ≥ 0.5:
        DMI_label ← "DMI+"
    ELSE:
        DMI_label ← "DMI-"
    STORE DMI_label
Step 6: Feature Fusion
FOR each patient:
image_features←extract_features_from(segmentation_map)
    combined_features ← CONCAT(image_features, clinical_features[patient])
    STORE combined_features
Step 7: Tumor Grading with ViT + LoRA
INITIALIZE VisionTransformer with LoRA_tuning
FOR each combined feature:
    patch_embeddings ← ViT.tokenize(combined_features)
    hidden_state ← ViT.forward(patch_embeddings)
    grade_prediction ← MLP_head(hidden_state)
    STORE grade_prediction
Step 8: Agentic Reasoning Module
FOR each prediction IN grade_prediction:
    reflection_result ← self_reflect(prediction, clinical_features)
external_knowledge_adjustment←consult_clinical_guidelines(reflection_result)
    final_grade ← revise_prediction(prediction, external_knowledge_adjustment)
    STORE final_grade
Step 9: Reporting by Agents
CALL EndometrialReportAgent with DMI_label, final_grade
CALL EndometrialVisionAgent to overlay segmentation
CALL EndometrialComplianceAgent to audit access logs
CALL EndometrialSchedulerAgent to monitor new uploads
Step 10: Output Results
FOR each patient:
    GENERATE results and tumor grade
END
```

Table 4. List of research model's hyperparameters

Component	Hyperparameter	Value / Setting	Remarks
MedSAM	Input Image Size	224 × 224	Matches ViT input size after downsampling
	Encoder Backbone	Vision Transformer (Frozen)	Pretrained; not fine-tuned during segmentation
	Prompt Type	Bounding Box	Used to segment the tumor and uterus regions
UCLGA	Line Fitting Method	Ellipse Fitting	Generates the uterine cavity line from segmentation maps
ViT with LoRA	Patch Size	16 × 16	Standard for ViT-base
	Number of Transformer Layers	12	ViT-Base configuration
	Embedding Dimension	768	Hidden size of transformer layers
	Heads in Multi-head Attention	12	Consistent with ViT-Base
	Dropout Rate	0.1	Helps regularization
	LoRA Rank	4	Number of rank factors added to attention weights
	LoRA Alpha	16	Scaling factor for low-rank matrices
	Learning Rate	0.0001	Optimized using validation
	Optimizer	AdamW	Weight decay + adaptive learning
	Epochs	50	Number of fine-tuning iterations
	Batch Size	32	Adjusted based on GPU memory
MLP Classifier	Hidden Layers	2 (512, 256 neurons)	Lightweight architecture
	Activation Function	ReLU	For non-linearity
	Output Units	3 (G1, G2, G3)	Multi-class classification
	Loss Function	Cross-Entropy Loss	Suitable for multi-class classification

4. Experiment and Analysis

4.1. Experiment Setup

In this section, the experimentation and analysis of the proposed EC detection and staging model, ViT-LoRA, are presented. The developed research model was experimented with and evaluated utilizing the PYTHON 3.7.12 programming language. The experimentations are conducted on Google Colab Pro. The ViT model fine-tuned using LoRA was evaluated for grading the EC stage. As the primary aim of this research is to find the depth of the MI, the EC stages are mainly detected as per FIGO guidelines.

4.2. Evaluation Metrics

The performance evaluation of the proposed model was carried out by conventional classification parameters such as accuracy, specificity, precision, F1 score, and sensitivity.

Accuracy: The efficiency of a model is assessed by its accuracy, determined by a ratio of TP and TN relative to all previous predictions. The following equation (24) is applied to compute the accuracy.

$$Accuracy = TP + TN/TP + TN + FP + FN \tag{24}$$

Precision: The TP rate, indicating the percentage of correctly predicted positive cases among all predicted instances, was utilized to evaluate the precision of the model. Equation (25) is utilized to assess the precision of the model.

$$Precision = TP/TP + FP \tag{25}$$

F1-Score: It is a precision and recall's harmonic mean and enables a balanced assessment of the performance of a model. Equation (26) is applied to calculate the F1-score.

$$F1 = 2 \times Precision \times Recall/Precision + Recall \tag{26}$$

Specificity: It is the measure of the model to accurately identify negative outcomes and is assessed by calculating the ratio of correctly classified TN. This statistic evaluates the classifier's potential to accurately identify and categorize adverse occurrences. The specificity of the model was computed by using equation (27).

$$Specificity = TN/TN + FP \tag{27}$$

Sensitivity: It refers to the model's potential to accurately identify positive instances. It quantifies the proportion of true positive cases accurately recognized.

It can also be referred to as recall. The following equation (28) is applied to compute the sensitivity of the model.

$$Sensitivity = TP/TP + FN \tag{28}$$

True Positive (TP) indicates the total instances of images properly classified as EC. True Negative (TN) characterizes the total images correctly designated as not stated as EC. False Positive (FP) is the total count of non-cancerous images inaccurately classified as EC when they are not. False Negative (FN) characterizes the overall count of cancerous image instances incorrectly classified as non-cancerous [16-30].

Table 5 elaborates on the statistical framework used in the research, delineating the performance metrics, dataset division, and control measures that guaranteed the proper and reproducible methodology. The significance of the differences among the cases of FIGO Stage IA and IB is shown in Table 6, stating that the applied clinical variables did not reveal any significant differences (p > 0.05), thus underlining the database of conventional clinical indicators that the advanced imaging-based and AI-driven analysis is necessary for accurate staging.

Table 5. Statistical analysis and evaluation methods used in the study

Aspect	Method / Metric	Assumption Details	Purpose
Descriptive Statistics (Clinical Data)	Mean ± Standard Deviation	Continuous variables are assumed to be approximately normally distributed	Summarize patient demographics and clinical characteristics
Categorical Data Analysis	Frequency and Percentage	Independent observations	Describe stage, grade, and invasion distribution
Dataset Split	75% Training / 25% Testing	Random split; no overlap between sets	Prevent data leakage and ensure fair evaluation
Performance Metrics	Accuracy	Based on TP, TN, FP, and FN from the confusion matrix	Overall classification correctness
	Precision	TP / (TP + FP)	Measure false-positive control
	Sensitivity (Recall)	TP / (TP + FN)	Measure true-positive detection capability
	Specificity	TN / (TN + FP)	Measure true-negative identification
	F1-Score	Harmonic mean of precision and recall	Balanced performance evaluation
Statistical Significance Level	$p < 0.05$	Two-tailed tests	Determine meaningful differences between groups
Group Comparisons (Stage IA vs IB)	Inferential statistical tests	Independent samples; normality assumed	Assess clinical variable significance
Control Measures	Fixed preprocessing, identical splits, same metrics	Consistent pipeline across experiments	Ensure reproducibility and fairness
Comparative Analysis	Same metrics for all benchmark models	Uniform evaluation protocol	Enable a valid performance comparison

Table 6. Statistical significance analysis of clinical variables between FIGO stage IA and stage IB

Clinical Variable	Stage IA	Stage IB	Statistical Test	p-Value
Age (years, mean ± SD)	48.4 ± 8.9	61.8 ± 9.3	Two-tailed t-test	0.99
Tumor Grade Distribution	G1–G3	G1–G3	Chi-square test	0.54
Maximum Tumor Diameter (<3 cm / ≥3 cm)	354 / 115	90 / 193	Chi-square test	0.54
Myometrial Invasion Depth (<50% / ≥50%)	456 / 13	19 / 264	Chi-square test	0.74
Mixed Carcinoma Presence (Yes / No)	192 / 277	149 / 134	Chi-square test	0.15

Table 7. Mean ± standard deviation of clinical variables across FIGO stages

Clinical Variable	FIGO Stage IA (Mean ± SD)	FIGO Stage IB (Mean ± SD)
Age (years)	48.4 ± 8.9	61.8 ± 9.3
Tumor Grade (Numerical Encoding*)	1.34 ± 0.56	1.78 ± 0.71
Maximum Tumor Diameter (cm)	2.41 ± 0.82	3.67 ± 1.14
Myometrial Invasion Depth (%)	18.6 ± 9.4	67.2 ± 12.1

In addition to this, Table 7 provides a summary of central tendency and variability of important continuous variables using mean ± standard deviation, showing the trends in age, tumor characteristics, and myometrial invasion depth on the basis of the observable stages, even though these trends are statistically non-significant.

4.3. Performance Assessment

The performance analysis of the developed research model is assessed based on both training and test datasets. The performance of the model is assessed utilizing common classification metrics like accuracy, specificity, precision, F1 score, and sensitivity. Both these performance assessments are tabulated individually in the following Tables 8 and 9.

Table 8. Results of proposed model using training dataset

Parameters	Normal	Abnormal
Accuracy	98.73	97.88
Precision	97.86	97.62
F1-score	98.43	97.57
Specificity	99.26	98.92
Sensitivity	98.40	97.86

Table 8 depicts the performance of the developed model using the training dataset, revealing a strong classification capability for both normal and abnormal endometrial cancer cases. It exhibits 98.73% accuracy for normal cases and an accuracy of 97.88% for abnormal cases, hence, correctly classifying most of the inputs given to it. A precision of 97.86% (normal) and 97.62% (abnormal), which is low on the false-positive rates, indicates that most of the positive instances predicted are indeed true. The F1-score, which involves balancing the recall and precision functions, was found to be 98.43% for the normal class and 97.57% for the abnormal class, confirming that the model equally performs in the two classes with no bias. Specificity values of 99.26% for the normal candidate and 98.92% for the abnormal candidate further reinforce the model's strength in correctly identifying true negatives, thereby lowering the chance of labelling a healthy patient as having EC. The sensitivity or recall, or true positive rate, of 98.40% (normal) and 97.86% (abnormal) reflects the model's ability to detect respective true-positive data. The results demonstrate that the model holds robustness, reliability, and suitability for situations demanding high stakes, such as early detection and EC staging. Figure 8 depicts the graphical illustration of the results on training data.

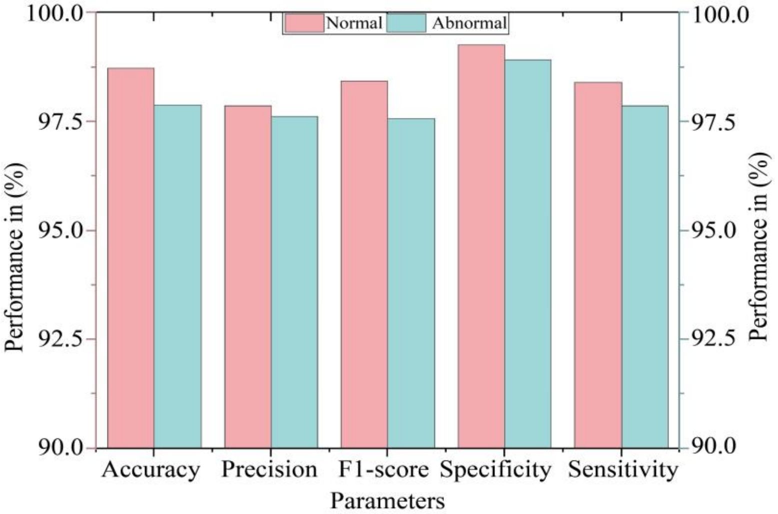


Fig. 8 Graphical illustration of results on training data

Table 9. Results of proposed model using test dataset

Parameters	Normal	Abnormal
Accuracy	97.59	96.82
Precision	97.25	96.77
F1-score	96.81	95.31
Specificity	98.74	96.64
Sensitivity	97.38	95.85

Table 9 shows the test performance of the developed model, consequently exhibiting its state-of-the-art generalizability and diagnostic capabilities in real-world environments. Accuracy values are 97.59% (normal) and 96.82% (abnormal), indicating that the model does not lose reliability in classification even on clinical data never seen before. The precision values, 97.25% (normal) and 96.77% (abnormal), prove that the model has a low false alarm rate, which is again extremely important in clinical applications, where maintaining reliability in reducing false alarms can be critical. The F1 values, which consider both precision and recall, remain at 96.81% (normal) and 95.31% (abnormal), which shows stability and balance for the performance of both classes. The specificity values for normal and abnormal cases are 98.74% and 96.64%. This shows the model's high potential to meaningfully control the false-positive rate and correctly declare the non-cancerous cases, avoiding unnecessary interventions or overdiagnosis. Since the reported sensitivity scores were 97.38 percent for normal cases and 95.85 percent for abnormal cases, the model has shown a strong ability to detect true positives, thus missing fewer cases of actual EC. Hence, it is evident from the results that the model offers a steady and clinically dependable performance and is ready for deployment as a tool for the detection and grading of preoperative EC in diverse patient populations. Figure 9 depicts the graphical illustration of the results on the test data.

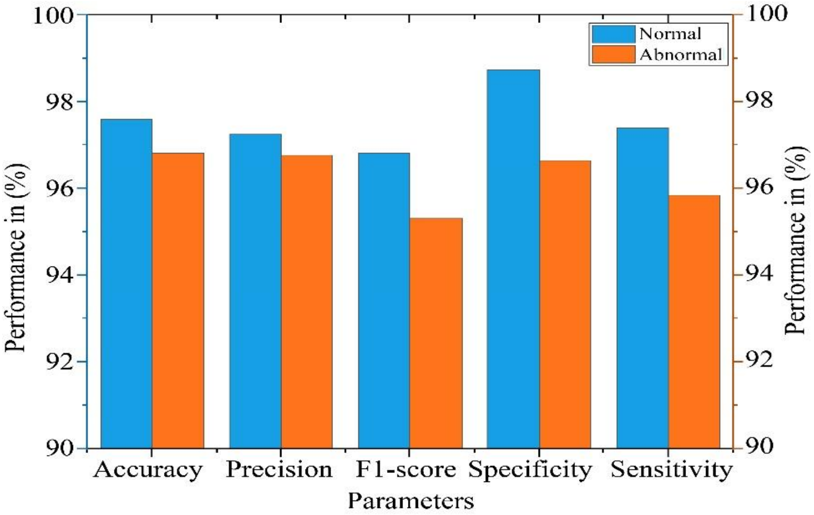


Fig. 9 Graphical illustration of results on test data

The results of the ablation study in Table 10 highlight the incremental contribution of each component in the MedSAM-ViT-LoRA framework and confirm the efficiency of the integrated design. The base CNN shows a lower performance in comparison with others, which indicates the drawbacks of traditional architectures for the complicated tasks of EC staging. The backbone of ViT is introduced, which contributes to an overall accuracy of capturing long-range spatial dependencies, and the addition of anatomical segmentation based on MedSAM boosts the performance even more by allowing the exact localization of the uterus and tumor areas. With the UCLGA integration, sensitivity and F1-score are boosted, made possible by the precise measurement of MI depth, a very important clinical marker. With the implementation of LoRA-based fine-tuning, the transformer model is adapted more efficiently with the least amount of parameters to obtain the surplus benefits as well. The proposed final model, including the agentic reasoning module, turns out to be the most effective one according to all metrics, thus underlining the significance of the processes of iterative self-reflection, clinical knowledge integration, and multi-agent collaboration for the generation of strong, clear, and clinically trustworthy predictions. Figure 10 depicts the graphical illustration of the ablation study results comparison.

Table 10. Comparison of ablation study results

Model Variant	Accuracy	Sensitivity	Precision	F1-Score	Specificity
Baseline CNN	89.12	88.40	87.95	88.17	90.03
ViT Only	92.68	91.72	92.15	91.93	93.10
MedSAM + ViT	94.21	93.10	94.02	93.56	94.48
MedSAM + ViT + UCLGA	95.34	94.42	95.10	94.76	95.68
MedSAM + ViT + LoRA	95.71	94.90	95.88	95.38	96.02
MedSAM + ViT + UCLGA + LoRA	96.08	95.30	96.10	95.68	96.21
Proposed Full Model	96.82	95.85	96.77	95.31	96.64

Table 11. Comparison of results with current models

Models	Accuracy	Sensitivity	Precision	F1-Score	Specificity
SSD-UNet [16]	86.90	81.80	NA	NA	91.70
U-Net [17]	88.60	92.30	NA	NA	86.40
RF [18]	85.71	95.45	84.00	90.00	69.23
ResNet-101 [19]	92.60	91.80	91.72	92.00	93.58
DenseNet-12+MLR [20]	84.20	94.10	76.20	68.50	74.40
EPSVM [21]	93.70	94.70	81.80	87.80	93.30
LR [22]	85.30	90.30	NA	NA	80.00

CNN [23]	89.00	93.00	NA	NA	70.00
Multi-sequence MRI Analysis [24]	90.30	91.70	84.60	NA	89.50
ResNet [25]	88.75	87.50	86.25	NA	95.00
Swin Transformer [26]	96.00	95.00	96.70	NA	96.70
mpMRI Analysis [27]	92.10	86.70	78.80	NA	93.60
MLP [28]	95.60	95.20	96.20	NA	96.90
NPM [29]	86.00	71.60	NA	NA	77.70
CNN [30]	91.00	94.00	NA	NA	91.00
MedSAM+ViT-LoRA	96.82	95.85	96.77	95.31	96.64

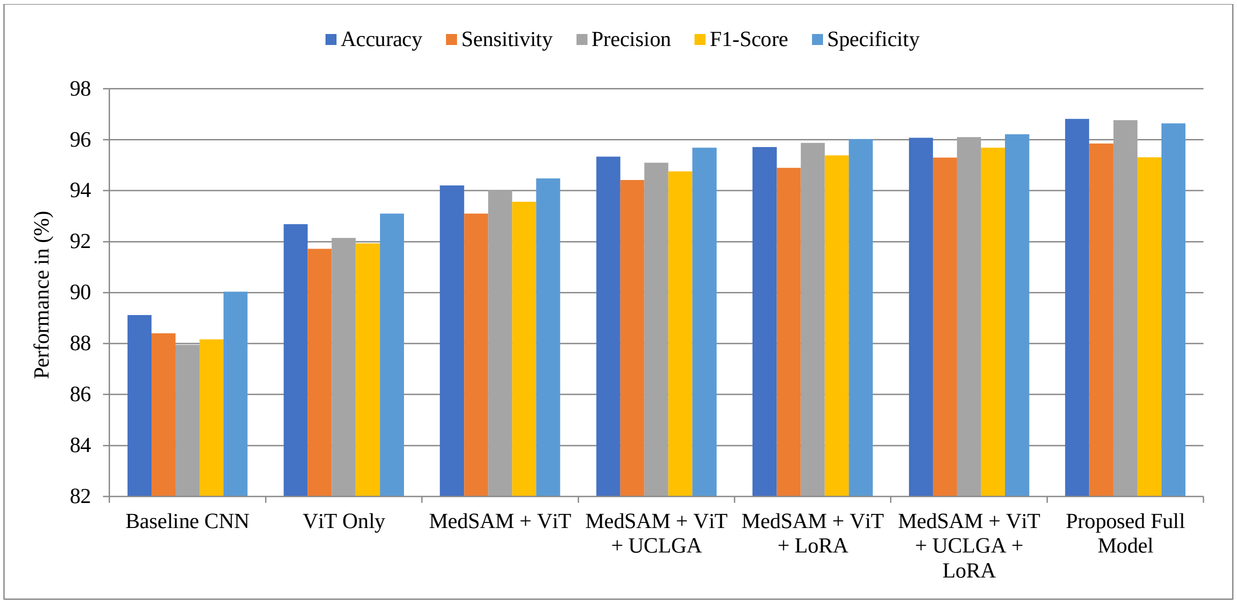


Fig. 10 Graphical illustration of ablation study results comparison

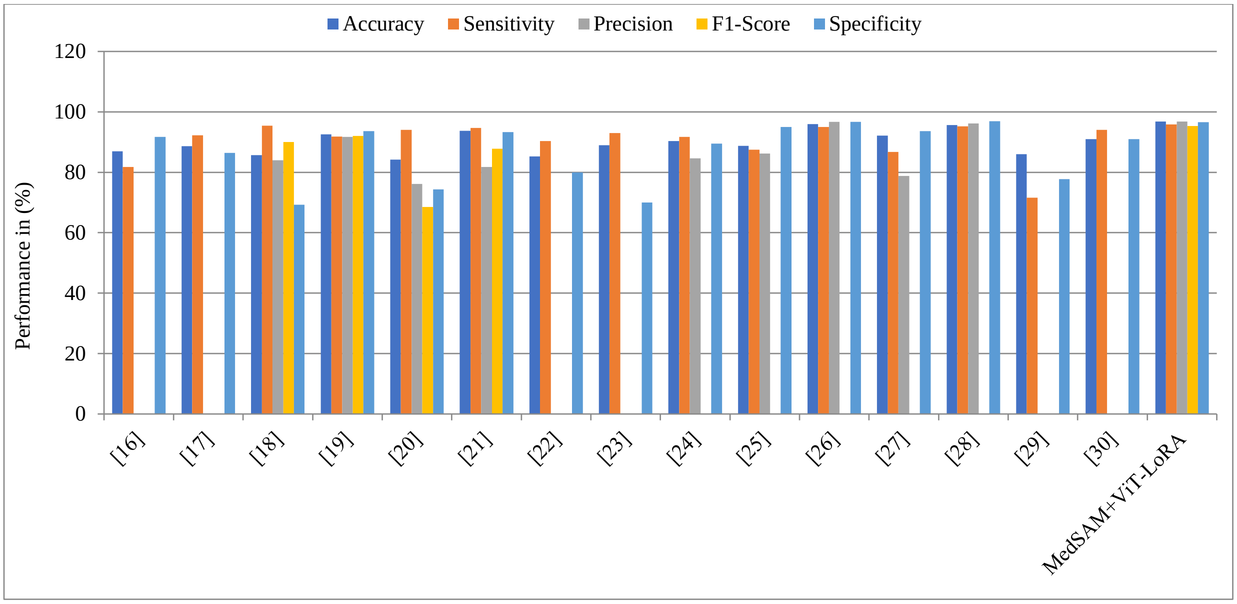


Fig. 11 Graphical illustration of results comparison

Table 11 presents a full comparison of the developed MedSAM+ViT-LoRA model with the state-of-the-art variations of EC detection and grading, exhibiting higher performance in all key evaluation metrics. The developed model yields the highest accuracy of 96.82 percent, surpassing the Swin Transformer with an accuracy of 96.00 percent, MLP with 95.60 percent, and EPSVM with 93.70 percent, thus confirming its strength in generalization and reliability. In terms of sensitivity (correctly identifying positive cases [abnormal]), the proposed model scores 95.85 percent, narrowly better than the Swin Transformer (95.00 percent) and ResNet-101 (91.80 percent), while markedly better than older methods, e.g., SSD-UNet (81.80 percent) and NPM (71.60 percent).

The precision of the developed model is 96.77%, indicating that fewer false positives exist in the model and that there is a higher confidence level in its positive predictions, which surpasses all prior models for which the value is present, e.g., ResNet-101 (91.72%) and RF (84.00%). Its F1-score, balancing precision and recall, is 95.31%, the highest among reported values and indicative of outstanding overall predictive harmony.

Then, with a specificity of 96.64%, the model performed well in identifying negatives (normal cases), trailing only MLP at 96.90% and ahead of EPSVM at 93.30% and U-Net at 86.40%. This consistent domination across all metrics underscores the strength of integrating advanced segmentation (MedSAM), transformer-based learning (ViT), and efficient fine-tuning (LoRA), alongside agentic reasoning; thus, positioning the developed model as the most all-inclusive and clinically viable solution among all compared approaches for highly accurate diagnosis of EC and tumor grading. Figure 11 depicts the graphical illustration of the results comparison.

The proposed framework outperformed existing state-of-the-art methods mainly because of the anatomically inspired and morphologically relevant design. In contrast to traditional methods that depend on localized characteristics or manually crafted descriptors like CNN and radiomics, the application of MedSAM allows for accurate segmentation of the uterus and tumor, so that the learning is very much concentrated on the areas of clinical significance.

The automated UCLGA gives a clear and precise measurement of MI, which is a major prognostic indicator frequently ignored in earlier studies. Moreover, the ViT model fine-tuned through LoRA is able to identify long-range spatial relationships while at the same time minimizing overfitting and the need for high computing power. An agentic reasoning module's incorporation results in decision-making consistency and interpretability improvement by means of iterative refinement and combining clinical expertise, thus providing a comprehensive explanation for the performance gains over traditional methods that were noticed.

4.4. Agentic-AI-based Grading Results

Figure 12 shows the interface of the Image Agent module. This module is part of the proposed Agentic AI framework for EC staging. It is an easy and intelligent MRI analysis tool that helps the clinician or researcher in the FIGO staging of endometrial cancer. The interface allows users to upload MRI scans either through drag-and-drop or click-to-upload in common file formats such as PNG, JPG, etc. After the MRI has been uploaded, the Image Agent analyzes the image and predicts the FIGO stage in an automated fashion by employing advanced AI models, such as MedSAM for segmentation and ViT-LoRA for feature extraction.

This module shows how the agentic system is designed with its users and clinical utility in mind while making radiological data-based staging decisions quickly, automatically, and accurately, aiding in timely diagnosis and treatment planning on a personalized basis.

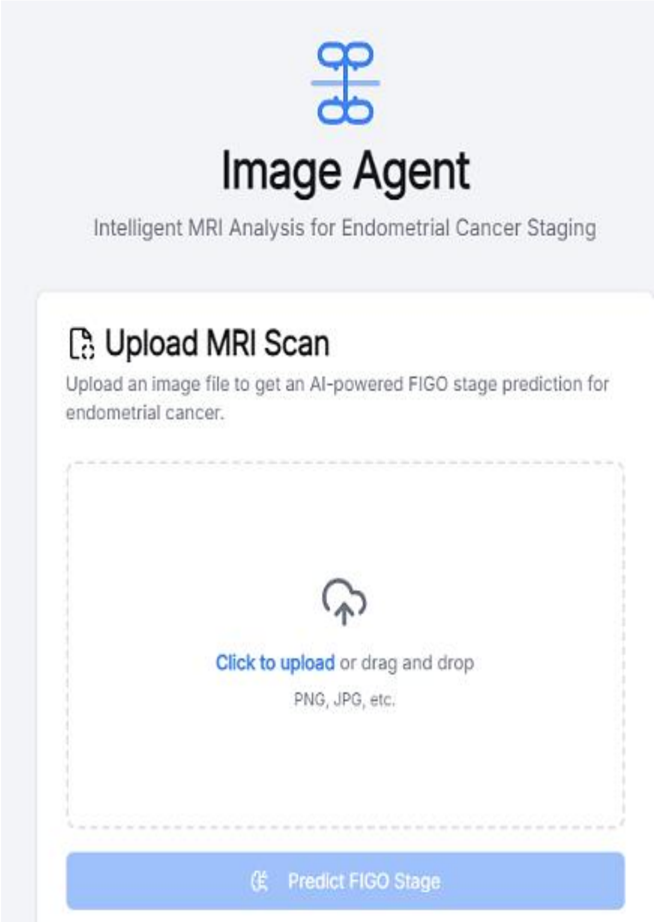


Image Agent

Intelligent MRI Analysis for Endometrial Cancer Staging

Upload MRI Scan

Upload an image file to get an AI-powered FIGO stage prediction for endometrial cancer.

Click to upload or drag and drop

PNG, JPG, etc.

Fig. 12 Image agent module interface

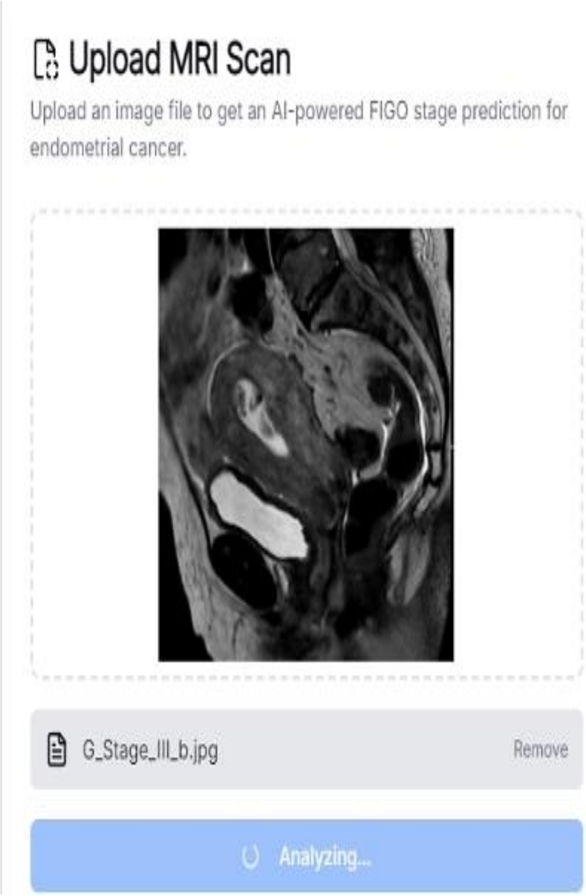
Figure 13 represents the active analysis phase of the Image Agent module within the Agentic AI system for EC staging. Here, an MRI image "G_Stage_III_b.jpg" has been uploaded, and now the system represents the "Analyzing..." state, which means the image is in the process of being handled in the AI pipeline.

The Image Agent, during this stage of analysis, triggers a downstream chain of operations such as segmentation (using MedSAM), identification of tumors and anatomical structures, and calculation of the depth of MI by UCLGA, while the ViT model, fine-tuned with LoRA, simultaneously extracts features needed for the prediction of the FIGO stage. Figure 14 shows the final output report for sample image 1, generated by the Image Agent module after analysis of the uploaded MRI scan, specifying the case as FIGO Stage IA of EC.

The Tumor Board Report section contains an overall clinical summary stating that myometrial invasion is less than 50%, with no evidence of cervical stroma or serosa or adnexa or lymph node or distant metastasis involvement scenario consistent with early-stage disease confined to the endometrium or inner half of the myometrium.

The system proposes a FIGO Stage IA diagnosis with 95% confidence, reflecting a high certainty in its prediction. The key findings gleaned from the report are myometrial invasion <50%, cervical stroma-No, and lymph node-No.

This output gave validation to the AI-based reasoner in interpreting MRI data and behaves similarly to a virtual tumor board when summarizing and stratifying patient cases to assist radiologists and oncologists in making evidence-based and time-sensitive decisions for early intervention and treatment planning.


Upload MRI Scan

Upload an image file to get an AI-powered FIGO stage prediction for endometrial cancer.

G_Stage_III_b.jpg

Fig. 13 Image agent module analysis stage

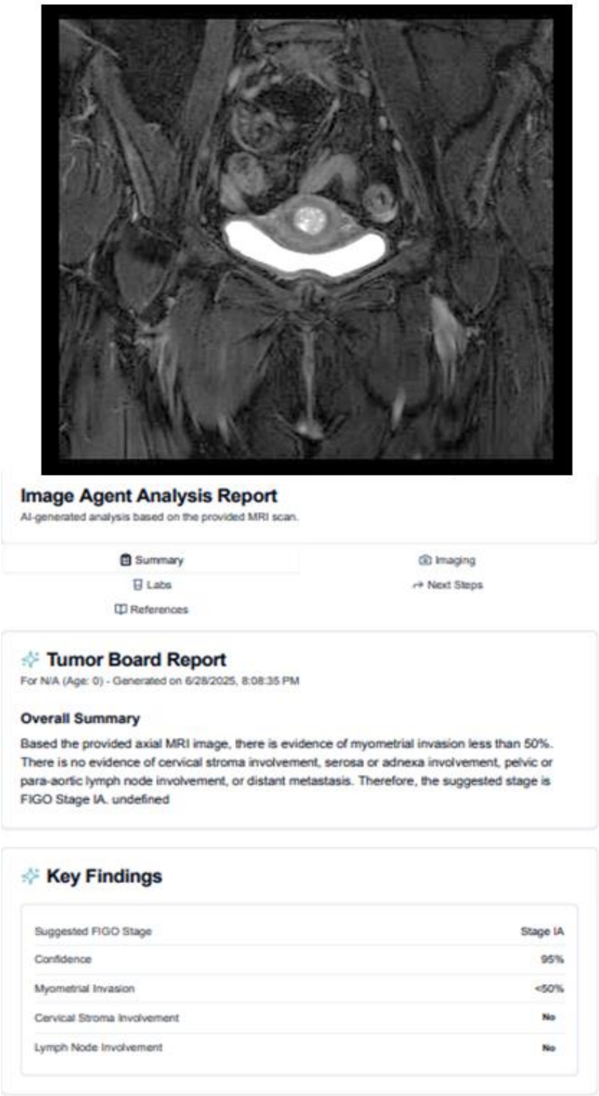


Image Agent Analysis Report

AI-generated analysis based on the provided MRI scan.

Tumor Board Report

For N/A (Age: 0) - Generated on 6/28/2025, 8:08:35 PM

Overall Summary

Based the provided axial MRI image, there is evidence of myometrial invasion less than 50%. There is no evidence of cervical stroma involvement, serosa or adnexa involvement, pelvic or para-aortic lymph node involvement, or distant metastasis. Therefore, the suggested stage is FIGO Stage IA. undefined

Key Findings

Suggested FIGO Stage	Stage IA
Confidence	95%
Myometrial Invasion	<50%
Cervical Stroma Involvement	No
Lymph Node Involvement	No

Fig. 14 Output report of sample image 1

Figure 15 shows the AI-generated Tumor Report of sample 2, designated as FIGO Stage II by Image Agent under Agentic AI. The Overall Summary states that more than 50% myometrial invasion and involvement of cervical stroma are seen on MRI, which are defining criteria for Stage II according to the FIGO classification system. It further states that no serosa or adnexal involvement is seen, with no lymph node or distant metastasis, thus discounting possible further progression to higher stages.

In support of this diagnosis and with a 95% confidence that indicates strong certainty of the model, the Key Findings state myometrial invasion ≥50%, cervical stroma involvement "Yes", and lymph node involvement "No". Such structured, clinically correct output demonstrates how the model uses image features to create accurate staging decisions and cements its role as a decision support system in preoperative planning for endometrial cancer management.

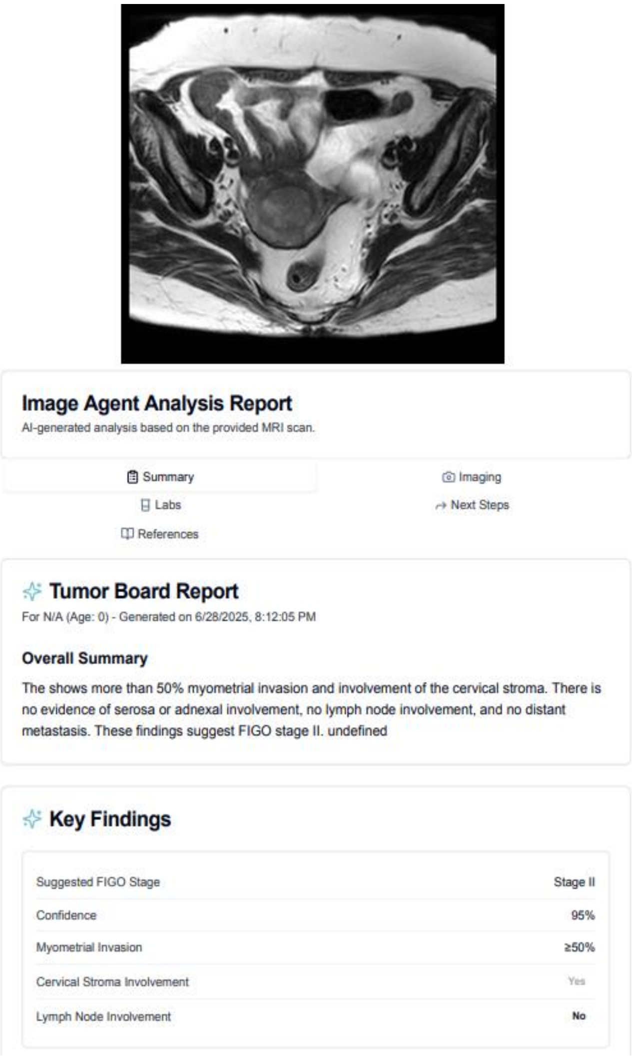


Image Agent Analysis Report

AI-generated analysis based on the provided MRI scan.

Tumor Board Report

For N/A (Age: 0) - Generated on 6/28/2025, 8:12:05 PM

Overall Summary

The shows more than 50% myometrial invasion and involvement of the cervical stroma. There is no evidence of serosa or adnexal involvement, no lymph node involvement, and no distant metastasis. These findings suggest FIGO stage II. undefined

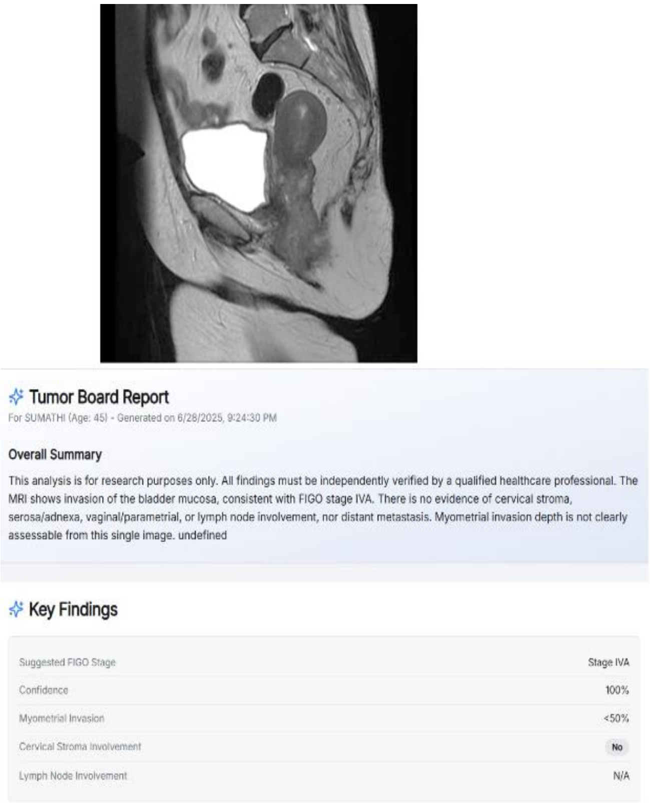
Key Findings

Suggested FIGO Stage	Stage II
Confidence	95%
Myometrial Invasion	≥50%
Cervical Stroma Involvement	Yes
Lymph Node Involvement	No

Fig. 15 Output report of sample image 2

Figure 16 shows the AI-generated Tumor Report for the patient identified under FIGO Stage IVA by the Image Agent within the Agentic AI framework. According to the Overall Summary, the MRI shows invasion of the bladder mucosa, which is the very hallmark of Stage IVA, in that direct extension of the tumor to adjacent pelvic organs takes place. The report finds no cervical stroma, serosa/adnexa, vaginal/parametrial, or lymph node involvement, along with the absence of distant metastasis.

Interestingly, generous information about myometrial invasion could not be gathered from this single image, while it is registered as "<50%" in the Key Findings; cervical stroma involvement is recorded as "No"; lymph node involvement is recorded as "N/A" based on the evidence being insufficient. The AI model gives 100% confidence in the assigned stage because of the high certainty in the interpretation of bladder wall invasion. This is an achievement for the model to handle complex staging decisions apart from offering aid in early detection and diagnosis of advanced-stage endometrial cancer.



Tumor Board Report

For SUMATHI (Age: 45) - Generated on 6/28/2025, 9:24:30 PM

Overall Summary

This analysis is for research purposes only. All findings must be independently verified by a qualified healthcare professional. The MRI shows invasion of the bladder mucosa, consistent with FIGO stage IVA. There is no evidence of cervical stroma, serosa/adnexa, vaginal/parametrial, or lymph node involvement, nor distant metastasis. Myometrial invasion depth is not clearly assessable from this single image. undefined

Key Findings

Suggested FIGO Stage	Stage IVA
Confidence	100%
Myometrial Invasion	<50%
Cervical Stroma Involvement	No
Lymph Node Involvement	N/A

Fig. 16 Output report of sample image 3

4.5. Advantages and Limitations

The developed research model includes numerous unprecedented advantages by way of an agentic AI framework that simultaneously considers MRI imaging and clinical data for correct EC staging and tumor grading. Consider the potential of segmentation through the utilization of MedSAM; uterine cavity line generation with UCLGA; and visual feature learning utilizing a ViT fine-tuned with LoRA, so that the model performs better classification and outperforms all of the metrics used by the state-of-the-art models.

The Agentic AI agent-based architecture, comprising Image Agent, Report Agent, Compliance Agent, and Scheduler Agent, facilitates automation, transparency, explainability, and clinical relevance. Meanwhile, the agentic reasoning module provides self-reflective capabilities along with the infusion of external knowledge, thereby improving the robustness of predictions and their contextual understanding.

However, the main limitation of the research is that it is related to the quality and diversity of MRI data, with some staging criteria perhaps not being fully assessable from single imaging slices (e.g., lymph node metastasis or distant spread). While recommendations generated via AI are performing best, they require clinical validation before implementation in a real-world scenario. The framework has laid a solid base for smart, explainable, and agent-driven cancer diagnostics.

5. Conclusion

A novel agentic AI approach performing automated staging and grading of EC was developed in this research, based on MRI imaging and clinical data within an end-to-end deep learning architecture.

The proposed system utilized MedSAM for accurate medical image segmentation, UCLGA for measuring myometrial invasion, and a ViT model fine-tuned with LoRA for deep feature extraction to mark significant anatomical and pathological markers. The agentic reasoning module acts to create interpretability and robustness by self-reflecting and incorporating relevant external clinical knowledge. The experimental results, along with the ablation study, confirm that all the modules have a significant role in enhancing the three aspects of accuracy, robustness, and interpretability.

It also reduced the requirements of manual input and inter-observer variability. The findings proved the model's superiority, wherein it reached the scores of 96.82% for accuracy, 95.85% for sensitivity, 96.77% for precision, 95.31% for F1-score, and 96.64% for specificity. The model was more effective than the recent methods used for the early detection of EC. In addition to this, the system can produce outputs in tumor board reports, which comprise FIGO staging, prognostic biomarkers, and confidence scores to help doctors in making timely decisions concerning treatment policies. Therefore, this study marks a milestone in AI-assisted gynecological oncology with the possibility of wider deployment of agentic AI systems in clinical trials of real-world workflows in clinical settings.

In the future, the research will focus on expanding the dataset to multi-institutional and multi-ethnic populations to improve generalizability, incorporating additional imaging modalities such as PET-CT for more comprehensive staging, and including longitudinal patient data to assist in predicting treatment outcomes. The study will primarily target large-scale multi-center validation, the incorporation of other imaging modalities and genome data, the creation of real-time clinical deployment tools, and the application of the agentic framework to personalized therapy response prediction, thereby extending the translational impact of the research. The further enhancement of the agentic reasoning module will consider clinician feedback in real time and validation to facilitate clinical acceptance and deployment in real-world diagnostic workflows.

Conflicts of Interest

This study was conducted in compliance with established ethical standards and guidelines for medical imaging research. The MRI dataset used in this research on EC was collected from the JSS Academy of Higher Education and Research and JSS Hospitals, Mysuru, Karnataka, India, after obtaining appropriate institutional approval. All patient data were anonymized prior to analysis to ensure privacy and confidentiality. As the study involved retrospective analysis of de-identified medical imaging data, informed consent was obtained or waived in accordance with the institutional ethics committee regulations. No personally identifiable information was accessed or disclosed at any stage of the study. The authors declare that there are no conflicts of interest associated with this research.

References

[1] Junmei He et al., “Accuracy of Radiomics in the Diagnosis and Preoperative High-Risk Assessment of Endometrial Cancer: A Systematic Review and Meta-Analysis,” *Frontiers in Oncology*, vol. 14, pp. 1-11, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Chenwen Sun et al., “Analysis of the Correlation between MRI Imaging Signs and Lymphovascular Space Invasion in Endometrial Cancer,” *Current Medical Imaging*, vol. 21, no. 1, pp. 1-10, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[3] Nabiha Midhat Ansari et al., “AI-Augmented Advances in the Diagnostic Approaches to Endometrial Cancer,” *Cancers*, vol. 17, no. 11, pp. 1-14, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[4] Gengshen Bai et al., “Artificial Intelligence Radiomics in the Diagnosis, Treatment, and Prognosis of Gynecological Cancer a Literature Review,” *Translational Cancers Research*, vol. 14, no. 4, pp. 2508-2532, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[5] Xuji Jiang et al., “Enhancing Clinical Decision-Making in Endometrial Cancer through Deep Learning Technology: A Review of Current Research,” *Digital Health*, vol. 10, pp. 1-12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Longyun Wang et al., “Diagnosis Test Accuracy of Artificial Intelligence for Endometrial Cancer Systematic Review and Meta-Analysis,” *Journal of Medical Internet Research*, vol. 27, pp. 1-15, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[7] István Madár et al., “Diagnostic Accuracy of Transvaginal Ultrasound and Magnetic Resonance Imaging for the Detection of Myometrial Infiltration in Endometrial Cancer: A Systematic Review and Meta-Analysis,” *Cancers*, vol. 16, no. 5, pp. 1-16, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Jing Yang et al., “MRI-based Intra Tumoral and Peritumoral Radiomic for Assessing Deep Myometrial Invasions in Patient with Early-Stages Endometrioid Adenocarcinoma,” *Frontier in Oncology*, vol. 14, pp. 1-11, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[9] Marco Gennarini et al., “Multi-Models Quantitative MRI of Uterine Cancer in Precision Medicine Era-A Narrative Review,” *Insight into Imaging*, vol. 16, no. 1, pp. 1-19, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[10] Vipul Bhardwaj et al., “Machine Learning for Endometrial Cancers Predictions and Prognostications,” *Frontier in Oncology*, vol. 12, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Diogo Miguel Machado Pereira et al., “The Updated 2023 Staging of Endometrial Cancers: Tips for MRI Interpretations,” *Abdominal Radiology*, vol. 51, no. 1, pp. 261-273, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[12] Octavia Petrila et al., “The Applicability of Artificial Intelligences in Predicting the Depths of Myometrial Invasions on MRI Studies-A Systematic Review,” *Diagnostics*, vol. 13, no. 15, pp. 1-10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Arnaldo Stanzione et al., “Training Radiology Resident to Evaluate Deep Myometrial Invasions in Endometrial Cancers Patient on MRI: A Learning Curve Study,” *European Journals of Radiology*, vol. 177, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Lise Lecointre et al., “Artificial Intelligence-Enhanced Magnetic Resonance Imaging-based Pre-Operative Staging in Patients with Endometrial Cancer,” *International Journal of Gynecological Cancers*, vol. 35, no. 1, pp. 1-7, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[15] Jawaria Changhez et al., “Evaluating the Efficacy and Accuracy of AI Assisted Diagnostic Techniques in Endometrial Carcinoma: A Systematic Review,” *Cureus*, vol. 16, no. 5, pp. 1-11, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Liu Xiong et al., “A Computer-Aided Determining Method for the Myometrial Infiltrations Depths of Early Endometrial Cancers on MRI Image,” *BioMedical Engineering OnLine*, vol. 22, no. 1, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Wei Mao et al., “A Deep Learning-based Automatics Staging Methods for Early Endometrial Cancers on MRI Image,” *Frontier in Physiology*, vol. 13, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Jing Wang et al., “A Predictions Model based on Deep Learning and Radiomic Feature of DWI for the Assessments of Microsatellites Instability in Endometrial Cancers,” *Cancer Medicines*, vol. 13, no. 16, pp. 1-13, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[19] Xinyu Qi, “Artificial Intelligences-Assisted Magnetic Resonances Imaging Technology in the Differential Diagnoses and Prognoses Predictions of Endometrial Cancers,” *Scientific Report*, vol. 14, no. 1, pp. 1-12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Jin Yang et al., “Combined Deep-Learning MRI-based Radiomics Model for Preoperative Risks Classifications of Endometrial Endometrioid Adenocarcinoma,” *Frontier in Oncology*, vol. 13, pp. 1-10, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Xueliang Zhu et al., “Detections of Deep Myometrial Invasions in Endometrial Cancers MR Imaging based on Multi-Features Fusions and Probabilistic Support Vector Machines Ensemble,” *Computer in Biology and Medicines*, vol. 134, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Qiu Bi et al., “Different Multiparametric MRI-based Radiomic Model for Differentiating Stages IA Endometrial Cancers from Benign Endometrial Lesion: A Multicenter Study,” *Frontier in Oncology*, vol. 12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Tsukasa Saida et al., “Differentiations of Carcinosarcoma from Endometrial Carcinoma on Magnetic Resonances Imaging using Deep Learning,” *Polish Journal of Radiology*, vol. 87, no. 1, pp. 521-529, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Sam Raja et al., “Enhancing Preoperative Assessments of Endometrial Cancers: The Roles of Diffusions-Weighted Magnetic Resonances Imaging in Evaluating Myometrial Invasions,” *Cureus*, vol. 16, no. 6, pp. 1-16, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Jingxiong Tao et al., “Evaluations and Monitoring of Endometrial Cancers based on Magnetic Resonances Imaging Feature of Deep Learning,” *Contrast Media and Molecular Imaging*, vol. 2022, no. 1, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[26] Ziyu Zheng et al., “Multimodal MRI Images Fusions for Early Automatic Staging of Endometrial Cancers,” *Sensors*, vol. 25, no. 9, pp. 1-16, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[27] Xuxu Meng et al., “Multiparametric Magnetic Resonances Imaging-based Assessments of the Effects of Adenomyosis on Determining the Depths of Myometrial Invasions in Endometrial Cancers,” *Quantitative Imaging in Medicines and Surgery*, vol. 14, no. 5, pp. 3717-3730, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[28] Huan Meng et al., “Predicting Risks Stratifications in Early-Stages Endometrial Carcinoma: Significances of Multiparametric MRI Radiomic Models,” *Journal of Imaging Informatics in Medicines*, vol. 37, no. 1, pp. 81-91, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[29] Dong-Hoon Jang et al., “Predictions of Final Pathology Depending on Preoperative Myometrial Invasions and Grades Assessments in Low-Risks Endometrial Cancers Patient: A Korean Gynecologic Oncology Group Ancillary Study,” *PloS One*, vol. 19, no. 6, pp. 1-17, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[30] Aiko Urushibara et al., “The Efficacy of Deep Learning Model in the Diagnoses of Endometrial Cancers using MRI: A Comparison with Radiologist,” *BMC Medical Imaging*, vol. 22, no. 1, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Yuan Chen et al., “Preoperative Discrimination of Absence or Presence of Myometrial Invasion in Endometrial Cancer with an MRI-based Multimodal Deep Learning Radiomics Model,” *Abdominal Radiology*, vol. 50, no. 7, pp. 3335-3346, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[32] Ran Guo et al., “Preoperative Prediction of Aggressive Endometrial Cancer using Multiparametric MRI-based Deep Transfer Learning Models,” *Frontiers in Oncology*, vol. 15, pp. 1-12, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[33] S. Manjunatha, and Malini M. Patil, “Interpolation Technique in Images Resampling,” *International Journal of Engineering and Technology*, vol. 7, no. 3.34, pp. 567-570, 2018. [Google Scholar]

[34] Iza Sazanita Isa et al., “Evaluating Denoising Performance of Fundamental Filter for T2-Weighted MRI Image,” *Procedia Computer Sciences*, vol. 60, pp. 760-768, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[35] Nema Salem, Hebatullah Malik, and Asmaa Shams, "Medical Images Enhancements based on Histograms Algorithm," *Procedia Computer Sciences*, vol. 163, pp. 300-311, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[36] Jun Ma et al., "Segment Anything in Medical Image," *Nature Communication*, vol. 15, no. 1, pp. 1-9, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[37] Christian Chang et al., "Segment Anything Models (SAM) and Medical SAM (MedSAM) for Lumbar Spines MRI," *Sensors*, vol. 25, no. 12, pp. 1-17, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[38] Zhenchen Hong et al., "Lightweight Low-Ranks Adaptations Vision Transformers Frameworks for Cervical Cancers Detections and Cervix Types Classifications," *Bioengineering*, vol. 11, no. 5, pp. 1-23, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[39] Chenqi Kong, Haoliang Li, and Shiqi Wang, "Enhancing General Faces Forgery Detections via Vision Transformers with Low-Ranks Adaptations," *2023 IEEE 6th International Conference on Multimedia Information Processing and Retrieval (MIPR)*, Singapore, pp. 102-107, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[40] Dyke Ferber et al., "Developments and Validations of an Autonomous Artificial Intelligences Agents for Clinical Decisions-Making in Oncology," *Nature Cancer*, vol. 6, no. 8, pp. 1337-1349, 2025. [CrossRef] [Google Scholar] [Publisher Link]