Journal of Cancer Research and Clinical Oncology (2023) 149:8877–8888
https://doi.org/10.1007/s00432-023-04838-4

RESEARCH

Deep learning-based methods for classification of microsatellite instability in endometrial cancer from HE-stained pathological images

Ying Zhang[1] · Shijie Chen[1] · Yuling Wang[1] · Jingjing Li[1] · Kai Xu[1] · Jyhcheng Chen[1] · Jie Zhao[1]

Received: 18 March 2023 / Accepted: 3 May 2023 / Published online: 7 May 2023

Abstract

Background Microsatellite instability (MSI) is one of the essential tumor biomarkers for cancer treatment and prognosis. The presence of more significant PD-L1 expression on the surface of tumor cells in endometrial cancer with MSI suggests that MSI may be a promising biomarker for anti-PD-1/PD-L1 immunotherapy. However, the conventional testing methods are labor-intensive and expensive for patients.

Methods Inspired by classifiers for MSI based on fast and low-cost deep-learning methods in previous investigations, a new architecture for MSI classification based on an attention module is proposed to extract features from pathological images. Especially, slide-level microsatellite status will be obtained by the bag of words method to aggregate probabilities predicted by the proposed model. The H&E-stained whole slide images (WSIs) from The Cancer Genome Atlas endometrial cohort are collected as the dataset. The performances of the proposed model were primarily evaluated by the area under the receiver-operating characteristic curve, accuracy, sensitivity, and F1-Score.

Results On the randomly divided test dataset, the proposed model achieved an accuracy of 0.80, a sensitivity of 0.857, a F1-Score of 0.826, and an AUROC of 0.799. We then visualize the results of the microsatellite status classification to capture more specific morphological features, helping pathologists better understand how deep learning performs the classification.

Conclusions This study implements the prediction of microsatellite status in endometrial cancer cases using deep-learning methods directly from H&E-stained WSIs. The proposed architecture can help the model capture more valuable features for classification. In contrast to current laboratory testing methods, the proposed model creates a more convenient screening tool for rapid automated testing for patients. This method can potentially be a clinical method for detecting the microsatellite status of endometrial cancer.

Keywords Microsatellite instability · Endometrial cancer · Whole slide image · Deep learning · Attention

Introduction

Mismatch repair deficiency can lead to the loss of expression of mismatch repair proteins, resulting in an unequal exchange of chromosomal or replication errors and base mismatches during DNA replication or recombination. This condition is also known as Microsatellite Instability (MSI) (Sinicrope and Sargent 2012). As an important genetic mechanism, it has been confirmed that MSI can cause Lynch Syndrome (LS), an inherited genetic disorder that can potentially cause cancer in multiple organs (Weiss et al. 2021). The incidence of MSI varies significantly among different types of cancers, with Endometrial Cancer (EC) having the highest probability of occurrence with MSI, reaching 31.4% (Bonneville et al. 2017). While the risk of EC in the general population is 3.1%, in patients with Lynch Syndrome, the risk of endometrial and colorectal cancer increases to 40–60% (Meyer et al. 2009). EC has become a significant threat to women's health and lives, accounting for 20–30% of gynecological diseases due to its increasing incidence (Bray et al. 2018).

The prognosis of EC is favorable for early stage tumors, while advanced cases with extrauterine metastases and high-risk histological types exhibit poor prognosis and low responsiveness to conventional treatments (Doll et al. 2014). Recent studies indicate that immunotherapy have shown

Ying Zhang and Shijie Chen have contributed equally to this work.

✉ Jie Zhao
zhaojie@xzhmu.edu.cn

[1] Xuzhou Medical University, Xuzhou 221004, Jiangsu, China

promising results in patients with recurrent and advanced EC associated with mismatch repair defects. The International Association of Gynaecologists also recommends MSI/dMMR testing for all EC patients (Concin et al. 2021). However, the conventional laboratory testing methods for MSI/dMMR, such as immunohistochemistry, polymerase chain reaction, and next-generation sequencing (Saeed et al. 2021), are expensive, require significant material and personnel resources, and may not be feasible for all patients. Deep learning methods have been used to detect microsatellite instability from pathological images in colorectal cancer with good results (Saillard et al. 2021; Echle et al. 2020; Lee et al. 2021). However, only a few studies have explored this approach in EC. The Manchester International Consensus Group also recommends that all EC patients should have their tumor tissue tested for dMMR/MSI status, regardless of the stage of the disease (Crosbie et al. 2019). Thus, a more convenient and less time-consuming testing method is needed to aid pathologists in the diagnosis and treatment of EC.

Pathological sections often reflect histopathological features that are linked to the molecular genetic information of the sample. Automated pathological examination can be performed on the Whole Slide Image (WSI), which can save time and reduce visual differences caused by human examination. In this study, WSIs obtained by scanning EC H&E-stained slides are used to predict microsatellite status directly from digital pathological images. At the same time, a Gated Recurrent Unit (GRU)-based Attention Module was added to the VGG16 network for improved classification performance. The attention module is added to measure how much importance the features contained in the image need to be given, so that the weights of different regions are correlated with the attention they receive, and thus, the key features that need to be attended are discovered. The processing steps were divided into data pre-processing, tissue classification, and microsatellite status classification. The output of previous step was filtered and used as input for the next process. An overview of the research is shown in Fig. 1, with details about the processing steps described below.

Related works

Deep learning

Deep learning (DL) has emerged as a prominent subfield of machine learning, particularly in computer vision and natural language processing. Unlike traditional machine learning, DL eliminates the need for complex data pre-processing by relying on deep neural networks that simulate the behavior of the human brain. These networks can receive and process unstructured data and automatically extract features, reducing dependence on human experts. Convolutional neural networks (CNNs) are a popular type of DL algorithm, with each convolutional layer extracting different information. Stacking multiple convolutional layers enables the network to extract increasingly complex and abstract features. Activation functions in the middle of convolutional layers enhance the network's ability to handle nonlinear problems and fit different distributions.

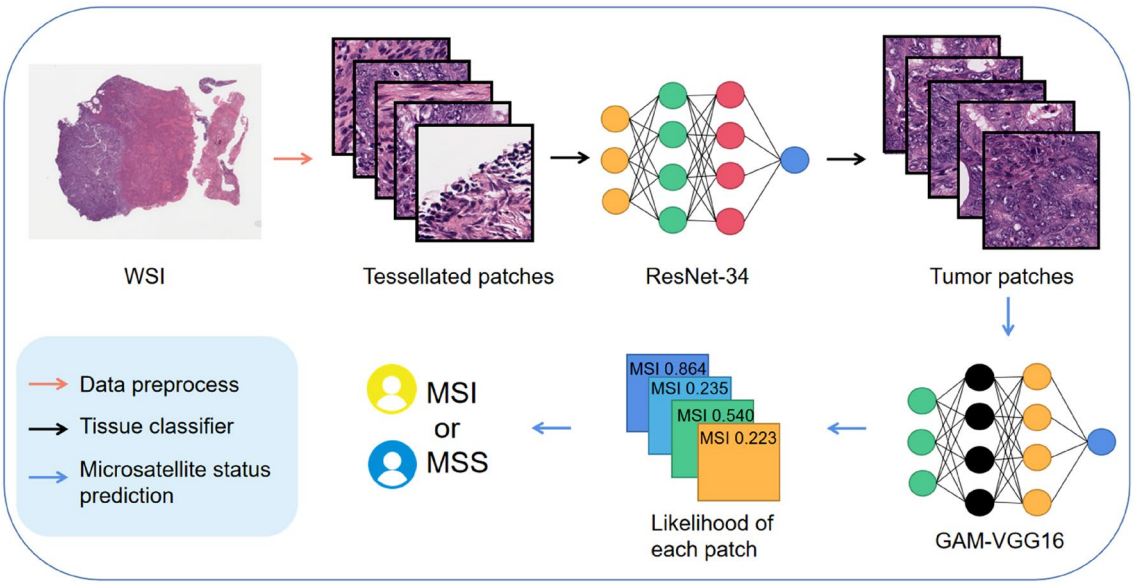

Fig. 1 An overview of studying deep learning-based classification of microsatellite instability in H&E-stained slide of endometrial cancer. *WSI* whole slide image, *MSS* microsatellite stability, *MSI* microsatellite instability

Classical network architectures, such as AlexNet (Krizhevsky et al. 2012), VGG (Simonyan and Zisserman 2015), and ResNet (He et al. 2016), have been proposed and widely used as backbones in more extended DL models. The availability of large-scale ImageNet databases in 2009 (Jia et al. 2009) has enabled researchers to train deeper and more stochastic networks; they are also using pre-trained networks to reduce training time.

Attention in deep learning

Inspired by human visual attention, researchers have applied visual selective attention models to DL so that neural networks can pay more attention to important local information and filter redundant noise information. Visual attention models can identify underlying correlations within the data and highlight significant features. Currently, the commonly used attention mechanisms include self-attention, spatial attention, channel attention.

Self-attention can learn the relationship between one pixel and all other position pixels, using the features of all bits to help generate a local detail in the image to highlight the effective features of this position. Transformer is the first model that entirely relies on the self-attention mechanism to compute the representation of its input and output, which has become a research hotspot in recent years. Spatial attention compresses digital feature information to prioritize spatial information. The Deepmind team proposed STL to achieve spatial invariance (Jaderberg et al. 2015). It does not require additional training supervision and can be added to the CNN to accomplish a specific spatial transformation of the input features. Channel attention mines the correlations in the convolutional channels and performs channel modeling to adjust the features channel by channel. SENet (Hu et al. 2020) models channel importance and enhances or suppresses feature communication for different tasks.

Deep learning with medicine

Over the past few years, there has been a growing adoption of DL techniques to support medical professionals in various aspects of disease diagnosis, treatment, and prognosis. Notably, image classification, segmentation, and reconstruction have gained significant traction as useful tools for clinical diagnosis and analysis (Astley et al. 2021). Song et al. (2019) established the benign and malignant classification model of thyroid nodules based on ultrasonic images, which adopted the Inception-V3 network architecture and classified the nodules by transfer learning. Zhao et al. (2017) proposed to integrate Fully Convolutional Networks (FCN) and Conditional Random Field (CRF) into a unified framework and developed a novel segmentation method for brain tumors. Besides classification and segmentation, detection is also the important area in medical artificial intelligence. Sapitri et al. (2023) proposed a YOLO-based method for the real-time detection of cardiac objects. This method used deep networks to detect the real-time fetal cardiac substructure with fetal ultrasound video. In a previous study (Zou et al. 2022), an easy-to-operate CNN with smaller parameters, namely TOD-CNN (Convolutional Neural Network for tiny object detection), was built for sperm detection; all of which have achieved good results.

Since the pathological examination is the gold standard for clinical cancer diagnosis, DL has paved the way for numerous exciting research topics in digital pathology. Many scholars have applied DL methods based on digital pathology to clinical diagnostic and therapeutic tasks, such as extraction of clinical biological structure information, including cellular level (Albarqouni et al. 2016), and glandular level (Fleming et al. 2012), disease grading, and prognosis, such as Gleason grading prostate cancer (Li et al. 2019), non-small cell lung cancer recurrence probability prediction (Coudray et al. 2017). Farajzadeh et al. (2022) introduced an end-to-end-based residual fully convolutional encoder–decoder network for localizing breast cancer on histopathology images. Also, DL with histopathological image can classify cancer grades. Dabass et al. (2022) proposed a network with multi-level convolutional and attention learning for classification of cancer grades and tissue structures in colon histopathological images. There have been several studies that utilize DL methods to identify molecular genetic features in digital pathological images, and these approaches have been shown to be more efficient compared to the conventional laboratory methods, resulting in significant savings of time and resources for patients.

Materials and methods

Dataset

This study involved the collection of WSIs of endometrial cancer from the public database known as The Cancer Genome Atlas (TCGA), where all the EC cases composed a cohort named TCGA-UCEC. The pyramid structure of a WSI is depicted in Fig. 2. TCGA is a comprehensive database that encompasses a wide range of genomic, epigenetic, and proteomic analyses across 33 different types of cancer, encompassing over 10,000 cases of tumors, aiming to leverage state-of-the-art genomic analysis techniques to gain a deeper understanding of the mechanisms underpinning cancer development and progression, with the ultimate goal of creating a complete genomic map of all human cancers. Two slide types are available in the TCGA-UCEC cohort, namely flash-frozen slides and Formalin-Fixed Paraffin-Embedded (FFPE) slides (Gabriel et al. 2013). The flash-frozen samples

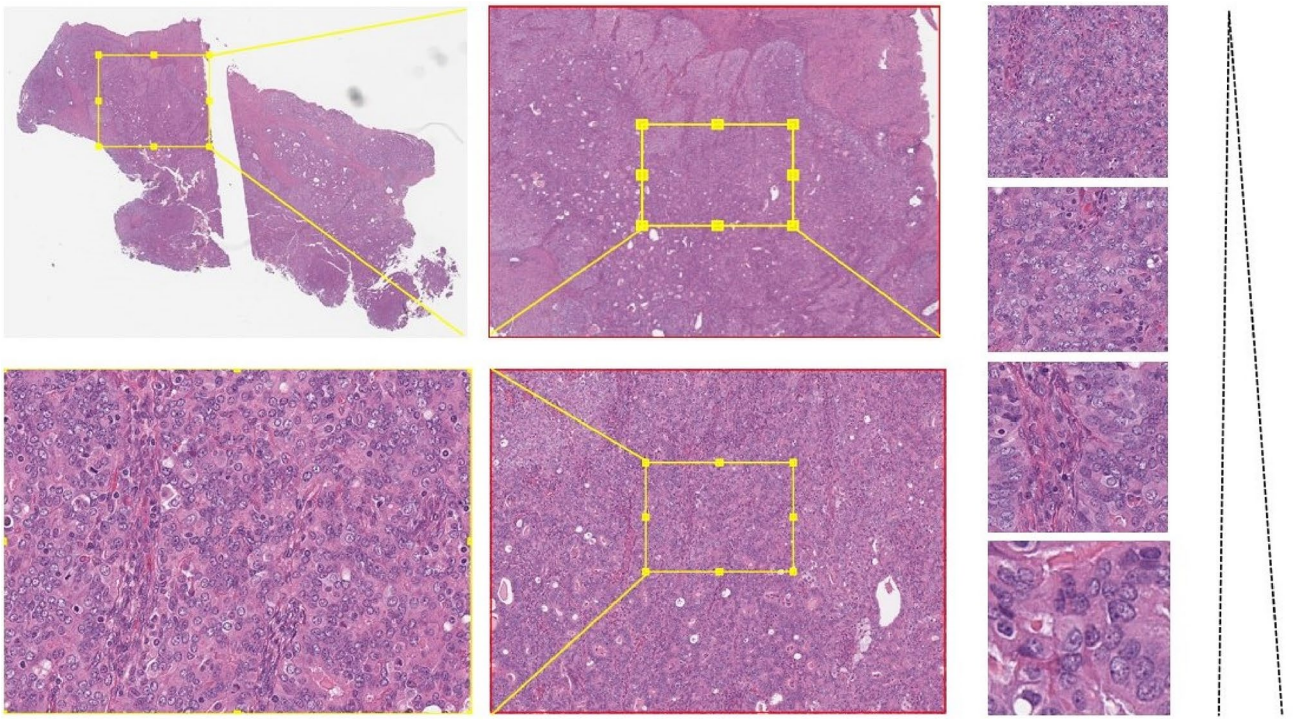

Fig. 2 The pyramid structure of a WSI. The section from overall appearance to local detail is shown on the left, and on the right, sampled images at different levels of resolution are shown

are typically produced during surgery to help the surgeon determine if the tumor has been fully resected, the quick process may potentially result in tissue damage. FFPE slides are produced through a series of distinct processes, resulting in a closer resemblance to actual tumor tissue. As a result, FFPE slides are considered a gold standard for diagnostic purposes. Following the initial screening of WSIs in the TCGA-UCEC cohort, the dataset used for this study was obtained. Here, we totally collected 95 FFPE WSIs from 95 patients, all WSIs can be downloaded on the TCGA websites, and the microsatellite status corresponding to the samples is provided in the auxiliary materials. ESMO guideline recommends that both MSI-L and MSS are classified as MSS (Luchini et al. 2019). A total of 47 MSI-H cases and 45 MSS cases were included in this study, with a balanced distribution between the two categories.

Methods

Data pre-processing and tissue classification

Due to the large resolution of the WSI, many operations on the image are restricted, making it unsuitable for direct input to CNNs. Thus, it is necessary to undertake appropriate pre-processing measures. First, we tessellate the different sizes of the WSI into tiles with a fixed shape of 256 × 256 pixels. Each WSI can be divided into an average of 16,000 small tiles, and then, we will discard tiles that contain limited or irrelevant information (see Fig. 3). The remaining tiles contain multiple tissue types, including fibers, fat, stroma, smooth muscle, and tumor epithelial tissue. To identify the histomorphological features associated with MSI, it is necessary to pay more attention to certain distinguishing features,

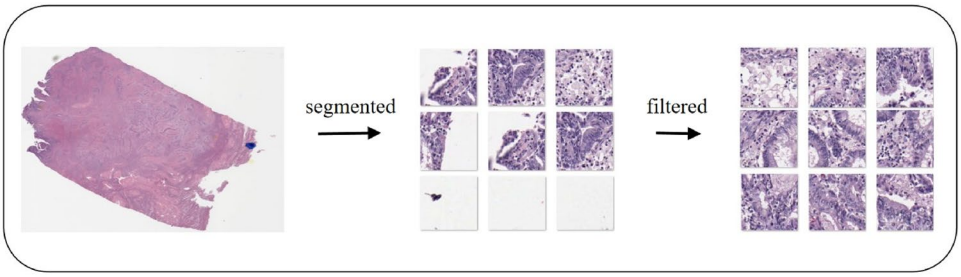

Fig. 3 An overview of the data pre-progress procedure

such as the presence of tumor-infiltrating lymphocytes and poorly differentiated tissue.

After obtaining the tiles that contain different tissues, the next step involves training a tissue classifier to categorize these tiles into three categories: ADIMUC, STRMUS, and TUMSTU (see Fig. 4). The classification method refers to Kather's (Kather et al. 2019a, b), where the ADIMUC category encompasses adipose and mucinous tissue, the STRMUS category includes stromal and muscle tissue, and the TUMSTU category represents tumor epithelial tissue. The training data named NCT_512_3CL is publicly accessible and freely available for download. The original dataset is comprised solely of patches extracted from colorectal cancer specimens. To improve the accuracy of tumor patch identification through tissue classification, a subset of tissue images from EC were added into the training dataset. The tissue classifier used the ResNet34 architecture pre-trained on ImageNet as the training model, with a batch size of 32. We utilized the cross-entropy loss function as the training strategy, which is defined as Eq. (1)

$$E = -\sum_{k} t_k \ \log \ y_k, \tag{1}$$

where log denotes the natural logarithm with e as the base, k denotes the index of dimensions of the data, y_k is the output of the neural network, and t_k is the true label of the correct solution. Additionally, 25% of the training data is split as a validation dataset to evaluate the tissue classifier and avoid overfitting.

Microsatellite status classifier

For the microsatellite status classifier, we used 70 WSIs for training and the remaining 25 cases were used as the test set. Importantly, the two datasets were non-overlapping, ensuring that the test set was entirely distinct from the training set. The delineation of the datasets is randomized, while adhering to a fixed random seed number for reproducibility purposes. The training data come from the red area of the tissue classification map (i.e., only the tumor area is preserved). We rely on the corresponding tissue map to tessellate the WSI at 20× magnification. In cases where the slides were scanned at 40× magnification, we increased the step size and subsequently resized the patch to ensure a consistent number of cells in the field of view as before. The DL-based workflow of the proposed GRU-based Attention Module (GAM) with VGG16 is shown in Fig. 5.

The microsatellite status prediction training set consists of 22,004 tumor patches that were randomly sampled from each type of MSS/MSI tissue, comprising of 12,002 MSS patches and 10,002 MSI patches. As part of the experimental setup, the VGG16 network is used to test baseline results on collected datasets and also as the backbone module combined with the GRU-Based Attention Module to compose the proposed architecture. GRU is a variant of the recurrent neural network (Kong et al. 2019). After removing the fully connected layer in the original VGG16 network, the feature maps extracted from the input data are subjected to an average pooling procedure, utilizing pooling kernels of varying sizes. Then, they are separately fed into the GAM after a simple reshape operation. GRU includes an update gate and reset gate (Dijk et al. 2016), which can extract more representative features by comparing spatial contextual information.

The architecture of our proposed GRU-Based Attention Module is shown in Fig. 6. The input size of module is B × HW × C. The initial hidden state is configured to have a size of 1 × B × C, while the GRU output size is B × HW × C. Then, two trailing fully connected layers aiming to obtain a

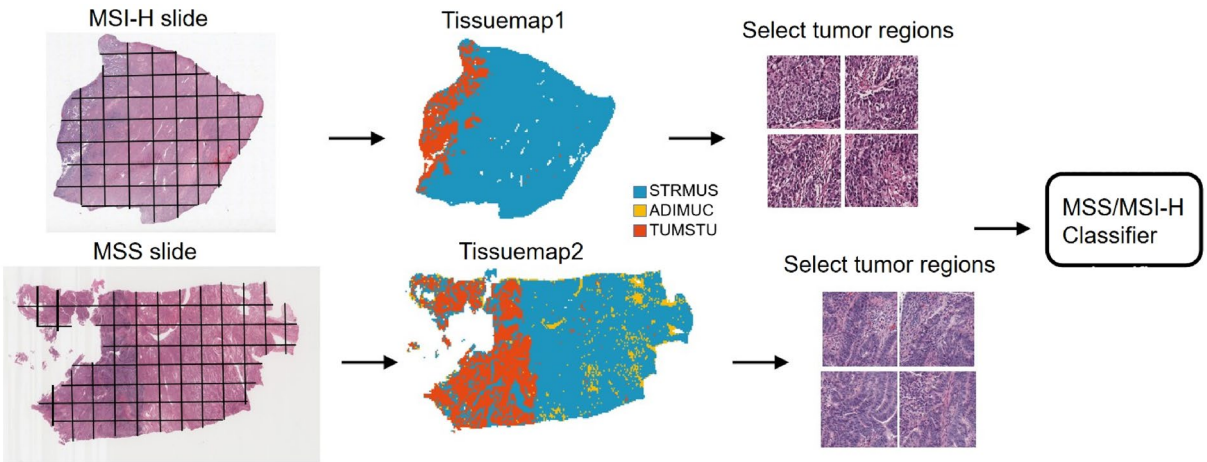

Fig. 4 The pipeline of tissue classification procedure

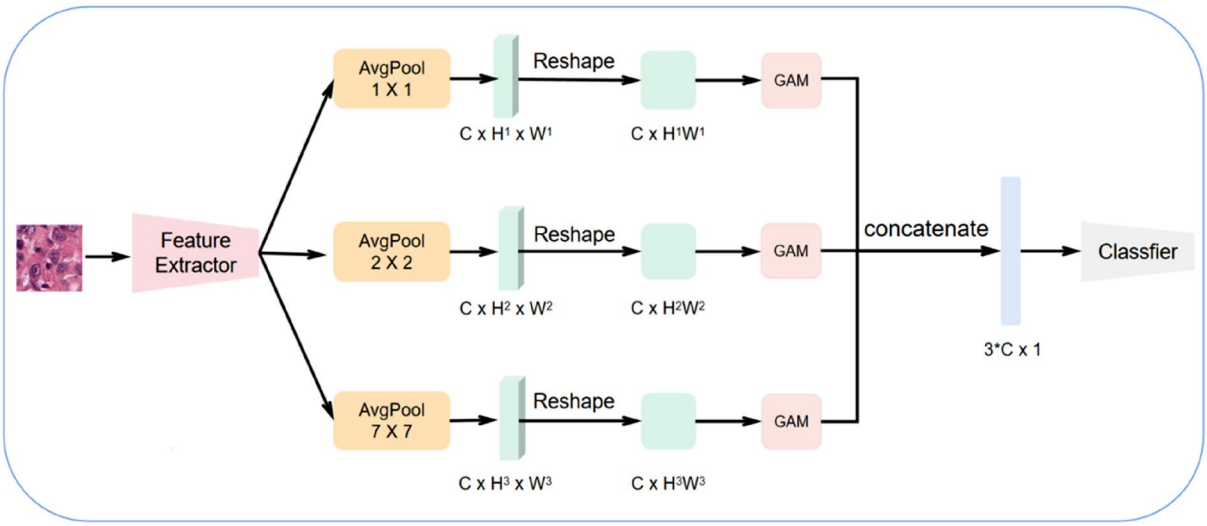

Fig. 5 The workflow of our proposed GAM-VGG16 model for microsatellite status classification

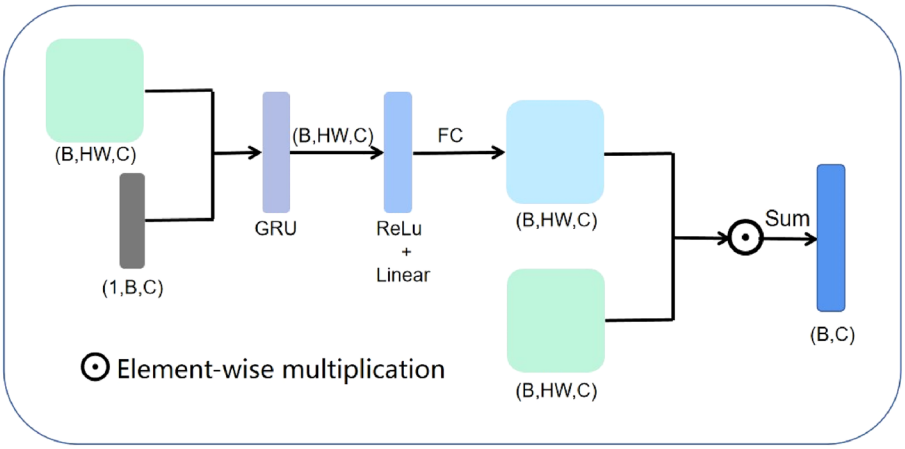

Fig. 6 The architecture of our proposed GRU-based attention module

feature layer with a size B × HW × C are set. By performing element-wise multiplication between the input layer and the output of shape B × HW × C, followed by summation along the first dimension, and the final feature layer with a size of B × C is obtained. The remainder consists of a concatenation calculation and a final SoftMax layer. We also used a cross-entropy loss function to calculate loss and used a stochastic gradient descent with a momentum of 0.9 and a weight decay of 1e−4 during training.

After predicting the MSI probability of a single patch, the microsatellite status of the entire slide needs to be determined. There are many ways to aggregate the predictions of all patches within slide level. Most scholars adopted straightforward monadic calculations such as calculating the mean probability (Yamashita et al. 2021). At the same time, there are also many complex methods, such as SVM, and some methods of multi-instance learning. Bag of Words (BoW) is a common document representation method in the field of information retrieval. In this process, each slide is treated as a document. Based on the patch-level likelihoods of WSIs, we can utilize the frequency patterns of terms to train a classifier for slide-level microsatellite instability prediction (Wang et al. 2020). We employ both the calculating mean probability method and the BoW method (Cao et al. 2020) for classification on slide level and compare the results of aggregation.

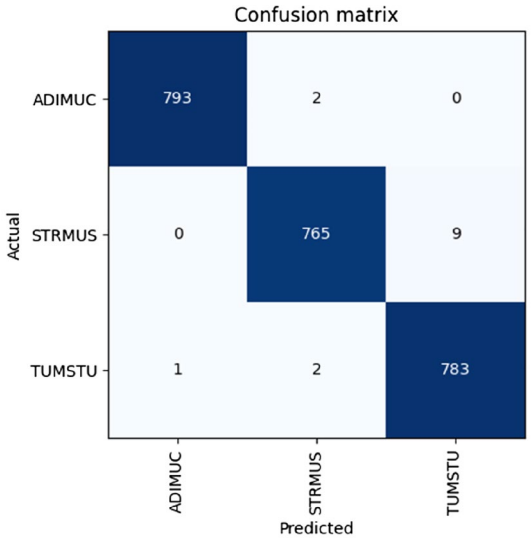

Fig. 7 Confusion matrix of the tissue classification. We randomly split 2355 images from our modified NCT_512_3CL dataset as validation data

Experiments and results

Performance evaluation of microsatellite status classifier

Our study aims to train a deep-learning model to predict microsatellite status based on WSIs. After applying the initial pre-progress implemented to the WSIs, the tessellated patches were subsequently fed into a convolutional neural network for tissue classification.

Pre-experiments' results

The modified NCT_512_3CL dataset achieved a validation accuracy of 0.9941 on a randomly divided validation set. The classification results of the validation set were depicted using a confusion matrix, which is presented in Fig. 7. Then, we used the saved model weights to make inferences on our dataset, and finally, we got the classification results of each tile. To visualize the classification results of the tissue classification model, a representative tissue map was generated by assigning different colors to tiles belonging to different tissue classes. The tissue maps are presented in Fig. 8. The red color represents the tumor epithelial tissue, loose non-tumor tissue is painted in blue

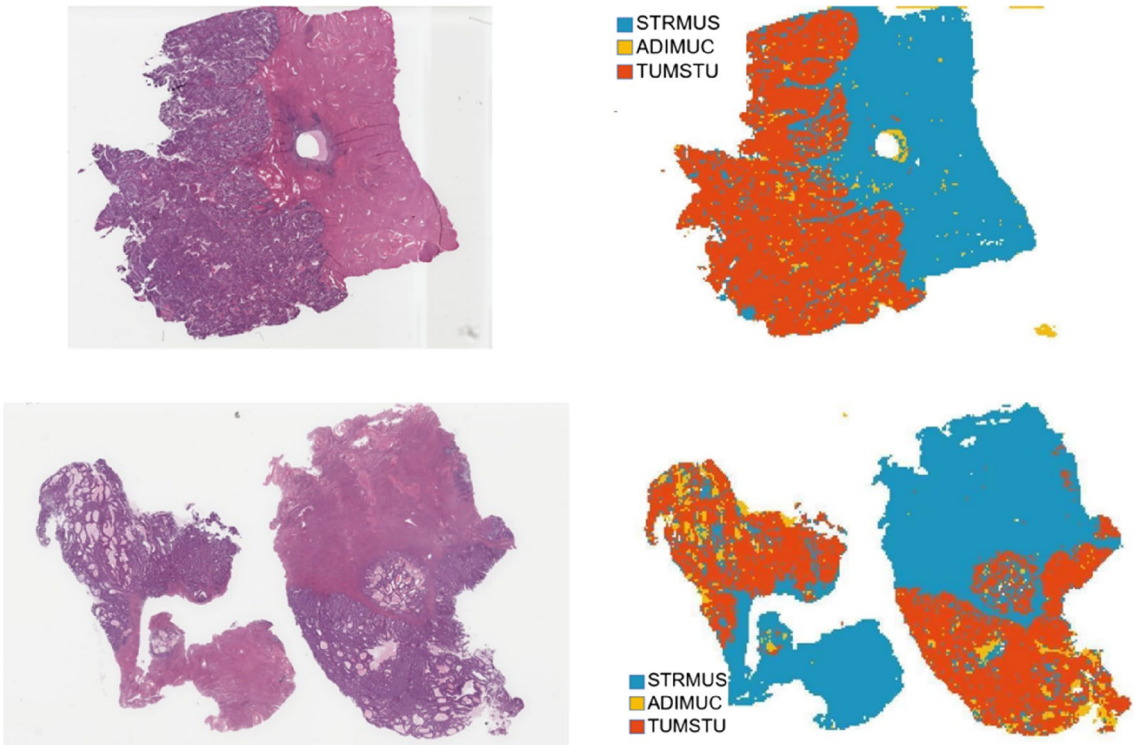

Fig. 8 Tissue classification for TCGA-UCEC cohort whole slide images. Each whole slide image corresponds to a tissue map for further operations

color, and the yellow color is dense non-tumor tissue. During the process of tessellating a WSI into small tiles, each individual tile is assigned corresponding coordinates that allow it to be mapped back to its original location within the WSI. This can provide valuable information for analyzing the sample and identifying any areas of interest tissue types. To assess the classification results obtained by the tissue classification model, we sought expert evaluation from pathologists. The pathologists provided positive feedback, stating that the classification boundaries were precise and that the model was effective in identifying tumor areas. The red area as the tumor epithelial tissue will be used as input into the MS status classifier for final slide-level status prediction.

As can be seen, most of the cancerous epithelial tissues are accurately assigned the red color. However, some misclassified tissue were observed, such as false predictions of some complex tumor tissues as normal cells, and it is speculated that the morphology heterogeneity of tumor cells may have influenced the classifier's judgment. Using neural networks to assist in automatically classifying tissues eliminates the need for the physician to manually outline tumor areas. This not only streamlines the classification process but also reduces the overall human involvement required, resulting in significant time savings.

Microsatellite classification results

The tissue classifier was successful in accurately identifying tumor region patches, resulting in an abundance of such patches. However, when training the MS classifier with whole number patches, the resulting classification performance was found to be mediocre. Thus, we implemented an undersampling approach. By reducing the number of tumor region patches in the training dataset, we aimed to achieve a more balanced distribution of MS training data. We first compared our GAM-VGG16 model with the baseline model (see Fig. 9A), and the prediction results can be observed that the improved VGG16-GAM model outperforms the baseline VGG16 model, demonstrating that the improvements we have made to the network are progressing. The baseline model with VGG16 achieved an AUC of 0.675, while our VGG16-GAM model improved to 0.799. The same parameters were set for the two comparison tests.

Other prediction results performed on the slide level are shown below. Here, the comparison tests were done in two aspects: (1) comparing the results before and after stain normalization (see Fig. 9B). In our study, we applied the Reinhard method as a stain normalization tool to minimize the impact of staining variability on the MS classification results (Reinhard et al. 2002). We found that the model was able to capture important features more accurately after applying the Reinhard method. (2) Comparing the results of the calculating mean probability method and

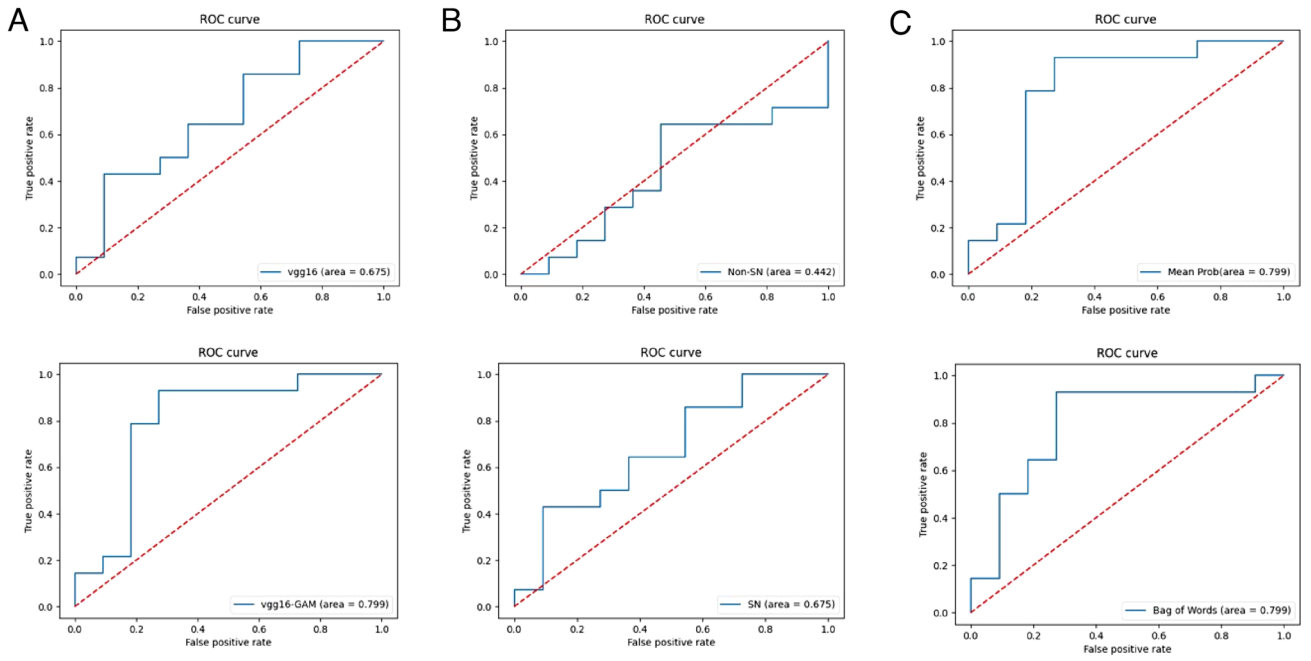

Fig. 9 **A** The comparison of our GAM-VGG16 model classification results with backbone. **B** The effect of stain normalization on classification results is compared based on otherwise identical training conditions. SN leads to significant improvement in classification performance. **C** The comparison between different aggregation methods

Table 1 When the aggregation methods were measured solely by AUC, the two methods were found to give comparable results, but BoW exhibits superior results when calculating other evaluation metrics, such as accuracy, precision, and F1-score

Option	Mean prob.	Bag of words
Accuracy	0.76	0.8
Precision	0.786	0.8
Recall	0.786	0.857
F1-score	0.786	0.826

BoW for slide-level prediction aggregation (see Fig. 9C). These two aggregation methods yield approximately the same AUC. However, the results show that BoW is superior to the calculating mean probability method when calculating other evaluation metrics, such as accuracy, precision, recall, and F1-score (see Table 1). The simple calculating method only considers the probabilities among the individual patches, which may lead to oversimplification of the information. In contrast, BoW method makes an average for the importance of different probabilities and pays more attention to the slide-specific patches.

Visualization of microsatellite status classification results

To enhance the interpretability of the MS status classifier and enable pathologists to better comprehend how neural networks arrive at conclusions in whole slide images, we have visualized the classification results. Generally speaking, the more pathologists trust the artificial intelligence (AI) models that assist in diagnosis, the more effective the assistance will be. There are various ways (e.g., probability heatmap, Grad-CAM) to increase the interpretability of the model, so that the model can better assist clinicians in their work.

We plotted the probability heatmap based on the prediction results, from which we can observe the visual features associated with MSI-H. To perform slide-level prediction, the WSI was tessellated into small patches due to its high resolution. After predicting small patches of the WSI using the convolutional neural network, these patches are given different prediction probabilities. To create a visual representation of the probability distribution, the patches are remapped onto the original whole slide image based on their location coordinates, as each patch has its corresponding probability value and, thus, a probability heatmap is

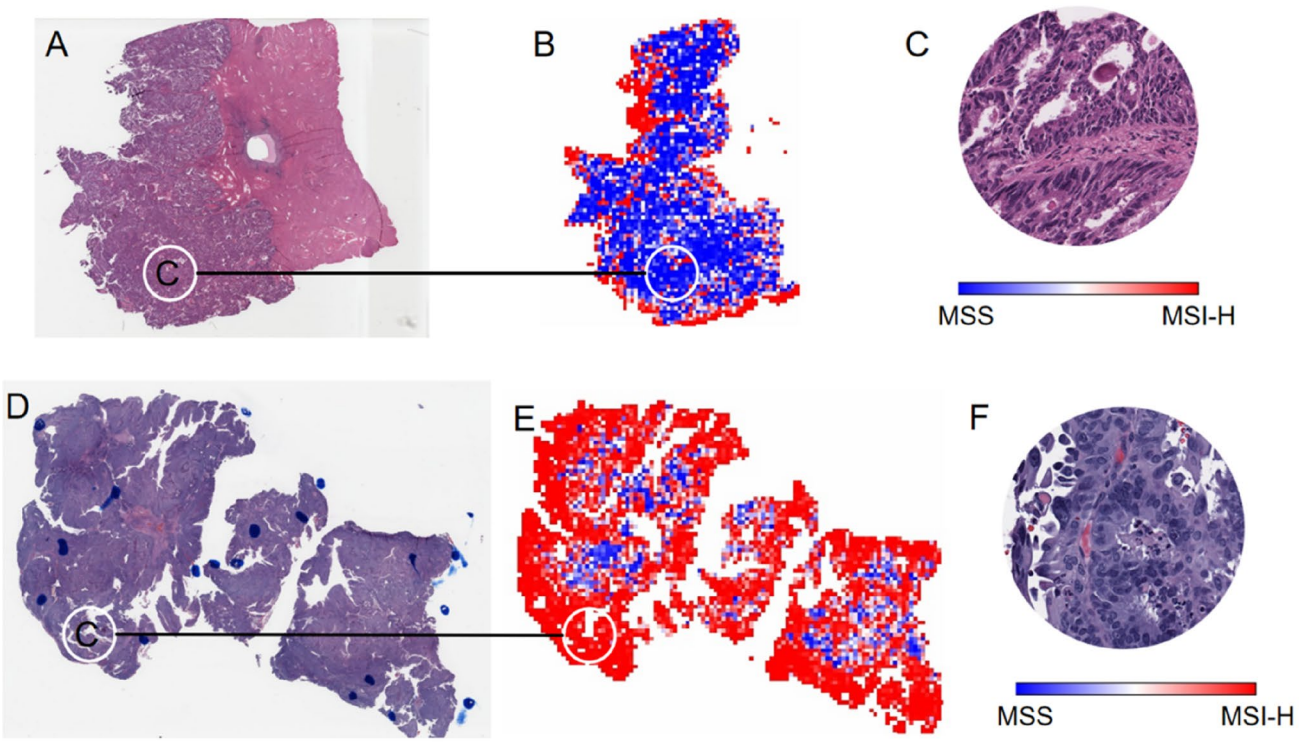

Fig. 10 The prediction map of two representative WSIs that were correctly predicted. Here, images are shown for each category, and they are in the order of the original whole slide image, the prediction map, and the image after partially enlarged. **A**–**C** Representative WSI of TCGA-UCEC cohort that is accurately classified as MSS. **D**–**F** Representative WSI of TCGA-UCEC cohort that is accurately classified as MSI-H

obtained (see Fig. 10). It can be observed that higher scored patches are given red color, indicating that those regions are more related to known MSI-H features, and patches with lower scores are less correlated with MSI-H patterns, which are colored blue. Zooming in on specific patches of tissue can indeed be useful for capturing and analyzing additional histomorphological features, which may be valuable for relevant histomorphological feature research.

Discussion and conclusion

With tumor biomarkers showing an increasingly important role in the treatment and prognosis of various cancers, more investigations have focused on implementing deep-learning approaches to construct faster and more accurate classifiers based on digital pathology for tumor biomarkers in clinical diagnostic and therapeutic tasks. In this study, FFPE slides with corresponding molecular features were collected from the endometrial cancer cohorts in the TCGA dataset. The slides underwent a pre-processing stage where they were segmented into smaller patches, so the patches can be used as input for the neural network during training. Given that the majority of WSIs contain a larger area of normal tissue than tumor tissue, deep-learning models have been utilized to classify various types of tissues. These models can effectively segment tumor epithelial cells during tissue classification processes, thereby alleviating the burden on pathologists who would otherwise have to perform this task manually. The utilization of isolated tumor regions within whole slide images can serve as a pivotal resource for tumor analysis research, as well as a potential asset for developing future machine learning models to extract distinctive image features.

Testing for microsatellite status is a valuable clinical tool that serves multiple purposes, including the diagnosis and screening for Lynch Syndrome, as well as providing insight into prognosis and immunotherapy prediction. We have presented a classification pipeline for detecting the microsatellite status with endometrial cancer. In the proposed classification model, the GAM-based model first predicts the patches individually and eventually integrates them to obtain the slide-level microsatellite status. GAM, as an attention module, successfully enables the architecture to capture more meaningful features from the feature maps generated by VGG16. Then, the aggregation approach of BoW is taken to obtain a prediction accuracy of 80% for the integrated slide-level status. Moreover, since the deep-learning methods have shown great potential in analyzing histopathological image analysis, our improved model can also be applied to other image classification tasks to achieve better performance.

There are still some limitations to our study that can perhaps be further explored. The results of the experiments may be slightly different due to the different processes of pathological specimens in different institutions, and different qualities of stained section specimens can impact classifier's performance (Cui and Zhang 2021). The positive impact of standardizing the steps in staining and preservation for improving the performance of the classification model needs more experiments to discuss and confirm. Although it has been reported that TCGA-UCEC cohort contains sources of cases from different regions and institutions, we may need to collect more recent cases in the future to verify the generalization of the proposed model. Noisy labels in the training data may affect classifier's performance; for example, cases classified as false positives in the test may be true cases of microsatellite instability which are missed by standard clinical tests. Some scholars have mentioned the problems of MSI test include: (1) 14% of patients failed the test because of low yield of DNA extraction or poor DNA quality. (2) When the purity of tumor cells in the test specimen is less than 30%, the result is likely to be false negative (Paula et al. 2021). Deep learning-based approaches outperform the conventional genetic or molecular tests in speed and convenience. However, they still require extensive testing and validation to ensure their safety and efficacy. In the future, we plan to collect larger scale clinical cases to further validate our model, which will enable more patients to benefit from the low-cost and general inspection method.

For patients with endometrial cancer, it is important to note that different tumor stages correspond to different prognoses, which makes clinical screening essential. A previous review has suggested that metabolomics may be useful for non-invasive screening, prediction of tumor histomorphological features, and tumor progression in EC patients (Raffone et al. 2020). Recently, an integrated machine learning approach appeared to validate this theory, and the results confirm that serum metabolism is indeed an accurate screening test for EC patients (Troisi et al. 2022). Meanwhile, many researchers have combined genomics with histopathology image data to make breakthroughs for advancing translational research in biomedical and computer-aided applications (Tan et al. 2021). Therefore, we can also try to contact pathological images with metabolomics for advancing our understanding of diseases and developing more effective treatment strategies, which may create extraordinary impacts on the diagnosis, treatment, and monitoring of endometrial cancer at the molecular and tissue level.

Acknowledgements The authors gratefully thank Doctor Li for her reviews and suggestions on pathology in this work. The authors also thank Xuzhou Medical University, The Affiliated Hospital of Xuzhou Medical University, and National Yang Ming Chiao Tung University for their support and help.

Author contributions YZ and SC made equal contributions to this work.

Funding This work is funded by the General Program of the China Postdoctoral Science Foundation under Grant No. 2019M651974.

Data availability The public datasets to support the results of this subject can be gained from TCGA (https://portal.gdc.cancer.gov/). The original training images for tissue classification are available at https://zenodo.org/record/2530789.

Declarations

Conflict of interest The authors declare that there is no conflict of interest regarding the publication of this paper.

References

Albarqouni S, Baur C, Achilles F et al (2016) AggNet: deep learning from crowds for mitosis detection in breast cancer histology images. IEEE Trans Med Imaging 35:1313–1321

Astley J, Wild J, Tahir B (2021) Deep learning in structural and functional lung image analysis. Br J Radiol 95:20201107

Bonneville R, Krook MA, Kautto EA et al (2017) Landscape of microsatellite instability across 39 cancer types. JCO Precis Oncol. https://doi.org/10.1200/po.17.00073

Bray F, Ferlay J et al (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68:394–424

Cao R, Yang F, Ma SC et al (2020) Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer. Theranostics 10:11080–11091. https://doi.org/10.7150/thno.49864

Concin N, Matias-Guiu X, Vergote I et al (2021) ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. Int J Gynecol Cancer 31:12–39. https://doi.org/10.1136/ijgc-2020-002230

Coudray N, Moreira AL, Sakellaropoulos T et al (2017) Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med 24:1559–1567

Crosbie EJ, Ryan N, Arends MJ et al (2019) The Manchester International Consensus Group recommendations for the management of gynecological cancers in Lynch syndrome. Genet Med 21:2390–2400

Cui M, Zhang DY (2021) Artificial intelligence and computational pathology. Lab Investig 101:412–422

Dabass M, Vashisth S, Vig R (2022) A convolution neural network with multi-level convolutional and attention learning for classification of cancer grades and tissue structures in colon histopathological images. Comput Biol Med 147:105680

Dijk B, Santos BF, Pita JP (2016) The recoverable robust stand allocation problem: a GRU airport case study. Or Spectrum 41:615–639

Doll KM, Tseng J, Denslow SA et al (2014) High-grade endometrial cancer: revisiting the impact of tumor size and location on outcomes. Gynecol Oncol 132:44–49. https://doi.org/10.1016/j.ygyno.2013.10.023

Echle A, Grabsch HI, Quirke P et al (2020) Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning. Gastroenterology 159:1406-1416.e11. https://doi.org/10.1053/j.gastro.2020.06.021

Farajzadeh N, Sadeghzadeh N, Hashemzadeh M (2022) A fully-convolutional residual encoder-decoder neural network to localize breast cancer on histopathology images. Comput Biol Med 147:105698

Fleming M, Ravula S, Tatishchev SF, Wang HL (2012) Colorectal carcinoma: pathologic aspects. J Gastrointest Oncol 3:153

Gabriel SB, Cibulskis K, Lander E et al (2013) Integrated genomic characterization of endometrial carcinoma. Nature 497:67–73

He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE conference on computer vision and pattern recognition (CVPR), pp 770–778

Hu J, Shen L, Albanie S et al (2020) Squeeze-and-excitation networks. IEEE Trans Pattern Anal Mach Intell 42:2011–2023

Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K (2015) Spatial transformer networks. MIT Press, Cambridge

Jia D, Wei D, Socher R et al (2009) ImageNet: a large-scale hierarchical image database, pp 248–255

Kather JN, Krisam J, Charoentong P et al (2019a) Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. PLoS Med 16:e1002730

Kather JN, Pearson AT, Halama N et al (2019b) Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med 25:1054–1056. https://doi.org/10.1038/s41591-019-0462-y

Kong J, Huang J, Yu H et al (2019) RNN-based default logic for route planning in urban environments. Neurocomputing 338:307–320

Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th international conference on neural information processing systems, vol 1. Curran Associates Inc., Red Hook, pp 1097–1105

Lee SH, Song IH, Jang HJ (2021) Feasibility of deep learning-based fully automated classification of microsatellite instability in tissue slides of colorectal cancer. Int J Cancer 149:728–740. https://doi.org/10.1002/ijc.33599

Li W, Li J, Sarma KV et al (2019) Path R-CNN for prostate cancer diagnosis and Gleason grading of histological images. IEEE Trans Med Imaging 38:945–954

Luchini C, Bibeau F, Ligtenberg MJL et al (2019) ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. Ann Oncol 30:1232–1243. https://doi.org/10.1093/annonc/mdz116

Meyer LA, Broaddus RR, Lu KH (2009) Endometrial cancer and lynch syndrome: clinical and pathologic considerations. Cancer Control 16:14–22. https://doi.org/10.1177/107327480901600103

Paula A, Delair DF, Ferrando L et al (2021) Genetic and molecular subtype heterogeneity in newly diagnosed early- and advanced-stage endometrial cancer. Gynecol Oncol 161:535–544

Raffone A, Troisi J, Boccia D et al (2020) Metabolomics in endometrial cancer diagnosis: a systematic review. Acta Obstet Gynecol Scand 99:1135–1146

Reinhard E, Ashikhmin M, Gooch B, Shirley P (2002) Color transfer between images. IEEE Comput Graph Appl 21:34–41

Saeed OAM, Mann SA, Luchini C et al (2021) Evaluating mismatch repair deficiency for solid tumor immunotherapy eligibility: immunohistochemistry versus microsatellite molecular testing. Hum Pathol 115:10–18

Saillard C, Dehaene O, Marchand T et al (2021) Self supervised learning improves dMMR/MSI detection from histology slides across multiple cancers, pp 1–16

Sapitri AI, Nurmaini S, Rachmatullah MN et al (2023) Deep learning-based real time detection for cardiac objects with fetal ultrasound video. Inform Med Unlocked 36:101150

Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition

Sinicrope FA, Sargent DJ (2012) Molecular pathways: microsatellite instability in colorectal cancer: prognostic, predictive, and therapeutic implications. Clin Cancer Res 18:1506–1512. https://doi.org/10.1158/1078-0432.CCR-11-1469

Song J, Chai YJ, Masuoka H et al (2019) Ultrasound image analysis using deep learning algorithm for the diagnosis of thyroid nodules. Medicine (baltimore) 98:e15133

Tan X, Su AT, Hajiabadi H et al (2021) Applying machine learning for integration of multi-modal genomics data and imaging data to quantify heterogeneity in tumour tissues. Methods Mol Biol 2190:209–228

Troisi J, Mollo A, Lombardi M et al (2022) The metabolomic approach for the screening of endometrial cancer: validation from a large cohort of women scheduled for gynecological surgery. Biomolecules. https://doi.org/10.3390/biom12091229

Wang T, Lu W, Yang F et al (2020) Microsatellite instability prediction of uterine corpus endometrial carcinoma based on H&E histology whole-slide imaging. In: IEEE international symposium on biomedical imaging, ISBI 2020

Weiss JM, Gupta S, Burke CA et al (2021) NCCN Guidelines®Insights: genetic/familial high-risk assessment: colorectal, version 1.2021. J Natl Compr Cancer Netw 19:1122–1132

Yamashita R, Long J, Longacre T et al (2021) Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. Lancet Oncol 22:132–141. https://doi.org/10.1016/S1470-2045(20)30535-0

Zhao X, Wu Y, Song G et al (2017) A deep learning model integrating FCNNs and CRFs for brain tumor segmentation. Med Image Anal 43:98–111

Zou S, Li C, Sun H et al (2022) TOD-CNN: an effective convolutional neural network for tiny object detection in sperm videos. Comput Biol Med 146:105543. https://doi.org/10.1016/j.compbiomed.2022.105543