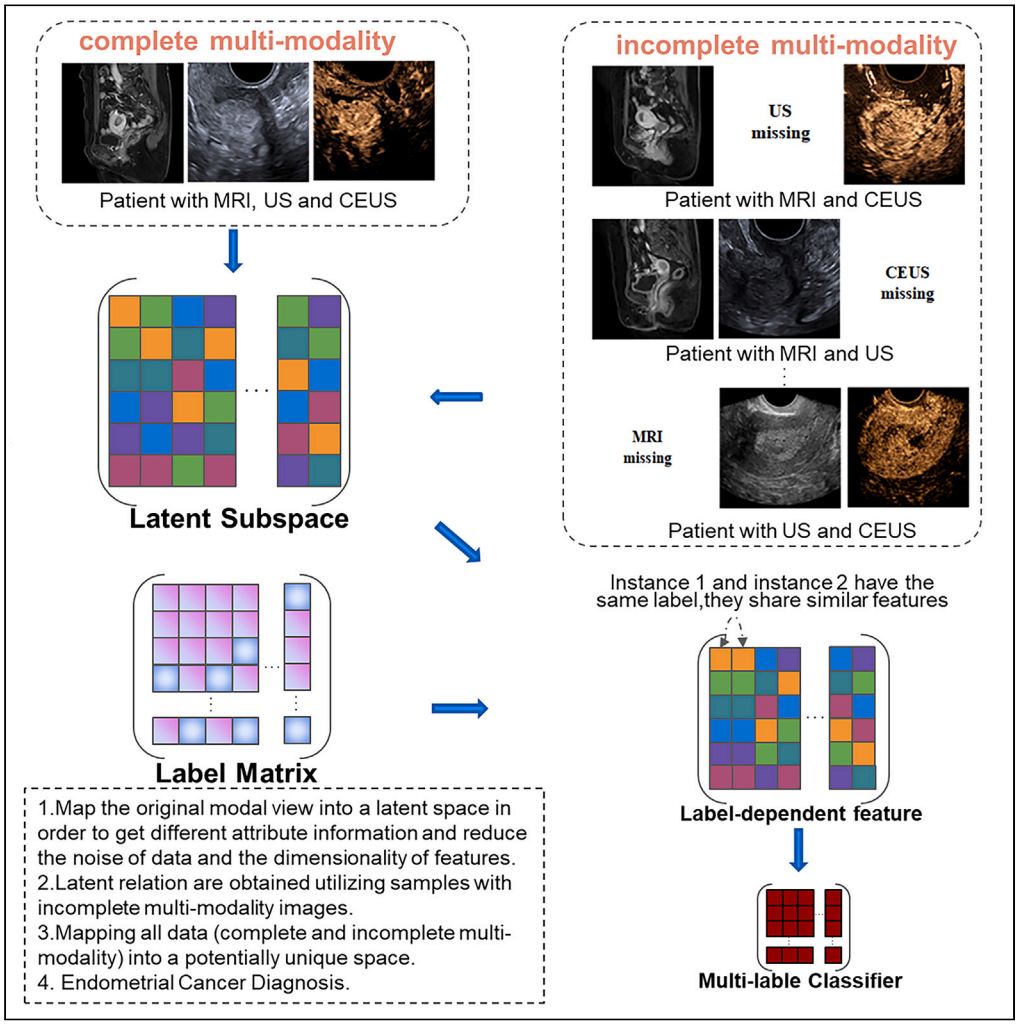
iScience

CellPress
OPEN ACCESS

Article

Latent relation shared learning for endometrial cancer diagnosis with incomplete multi-modality medical images



Jiaqi Li, Lejian Liao, Meihuizi Jia, Zhendong Chen, Xin Liu

liaolj@bit.edu.cn (L.L.)
xin3929@163.com (X.L.)

Highlights

An incomplete multi-modality medical images classification method is introduced

Latent relation shared learning boosts sensitivity in incomplete multimodality models

The fusion of MRI, US and CEUS methods enhances diagnostic performance

Li et al., iScience 27, 110509
August 16, 2024
https://doi.org/10.1016/j.isci.2024.110509

iScience

CellPress
OPEN ACCESS

Article

Latent relation shared learning for endometrial cancer diagnosis with incomplete multi-modality medical images

Jiaqi Li,[1,2] Lejian Liao,[1,2,4,*] Meihuizi Jia,[1,2] Zhendong Chen,[1,2] and Xin Liu[3,4,5,*]

SUMMARY

Magnetic resonance imaging (MRI), ultrasound (US), and contrast-enhanced ultrasound (CEUS) can provide different image data about uterus, which have been used in the preoperative assessment of endometrial cancer. In practice, not all the patients have complete multi-modality medical images due to the high cost or long examination period. Most of the existing methods need to perform data cleansing or discard samples with missing modalities, which will influence the performance of the model. In this work, we propose an incomplete multi-modality images data fusion method based on latent relation shared to overcome this limitation. The shared space contains the common latent feature representation and modality-specific latent feature representation from the complete and incomplete multi-modality data, which jointly exploits both consistent and complementary information among multiple images. The experimental results show that our method outperforms the current representative approaches in terms of classification accuracy, sensitivity, specificity, and area under curve (AUC). Furthermore, our method performs well under varying imaging missing rates.

INTRODUCTION

Endometrial cancer is the most common malignancy of the inner lining of the uterus and the fifth most common cancer in women, with 90000 deaths per year worldwide.[1,2] There are FIGO staging classification for endometrial cancer, and the treatment of patients varies with different clinical stages.[3] Preoperative assessment of myometrial/cervical invasion and lymph node metastases can help to define resulting surgical management. Magnetic resonance imaging (MRI) and transvaginal ultrasound are considered as the common imaging techniques for preoperative assessment of endometrial cancer. MRI has high diagnostic accuracy due to its excellent soft tissue contrast resolution.[4,5] However, MRI is not the first choice due to its high price and long queuing time. Ultrasound (US) is the primary screening method for endometrial cancer because of its low cost and reasonably sensitive and specific. Moreover, when the endometrial thickness is less than 5 mm, ultrasound is an effective first test with a high negative predictive value.[6]

Combining multimodal data can gain more information about the disease, and diagnose the disease more accurately.[7–9] For instance, multi-modality fusion learning architecture GroupFusionNet (GFN) method is used to diagnose theoptic neuropathy.[10] COVID-19 patient severity prediction is carried out through the cohesive multi-modality feature learning and fusion method.[11] Multi-level multi-modality (PET and CT) fusion radiomics method[12] have been used for the prediction of non-small cell lung carcinoma. Due to the high cost of some medical image examinations or the long queue time for examinations, etc., not all patients have complete multi-modality data.[13,14] Therefore, algorithms that require complete modal data are difficult to implement in the field of medical images. The missing modalities often contain a large amount of useful information, and missing some modal data in the limited data of a particular disease may lead to modeloverfitting. Ren[15] proposed a cross-site prognosis prediction framework for nasopharyngeal carcinoma from incomplete multi-modal data. Zhang[16] proposed a scalable swin transformer network using a single encoder to extract latent feature maps from incomplete modalities. So, not only do we need to mine data for potential relationships in existing modalities, but we should also make joint complements for missing parts. It is difficult to train some features appearing in only one or certain components of medical images, especially when the number of training data are limited,[17] it becomes very important to make full use of the existing medical image information.

In clinical work, doctors diagnose diseases by observing and analyzing the existing medical images of patients. It's a hard work, and takes years for an experienced doctor to read thousands of medical images to acquire a certain ability to diagnose diseases. At the same time, there may be slight subjective differences in the diagnosis of the same medical image by different doctors.[17]

[1]School of Computer Science and Technology, Beijing Institute of Technology, Beijing 100081, China
[2]Beijing Engineering Research Center of High Volume Language Information Processing and Cloud Computing Applications, Beijing 100081, China
[3]Department of Ultrasound, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China
[4]These authors contributed equally
[5]Lead contact
*Correspondence: liaolj@bit.edu.cn (L.L.), xin3929@163.com (X.L.)
https://doi.org/10.1016/j.isci.2024.110509

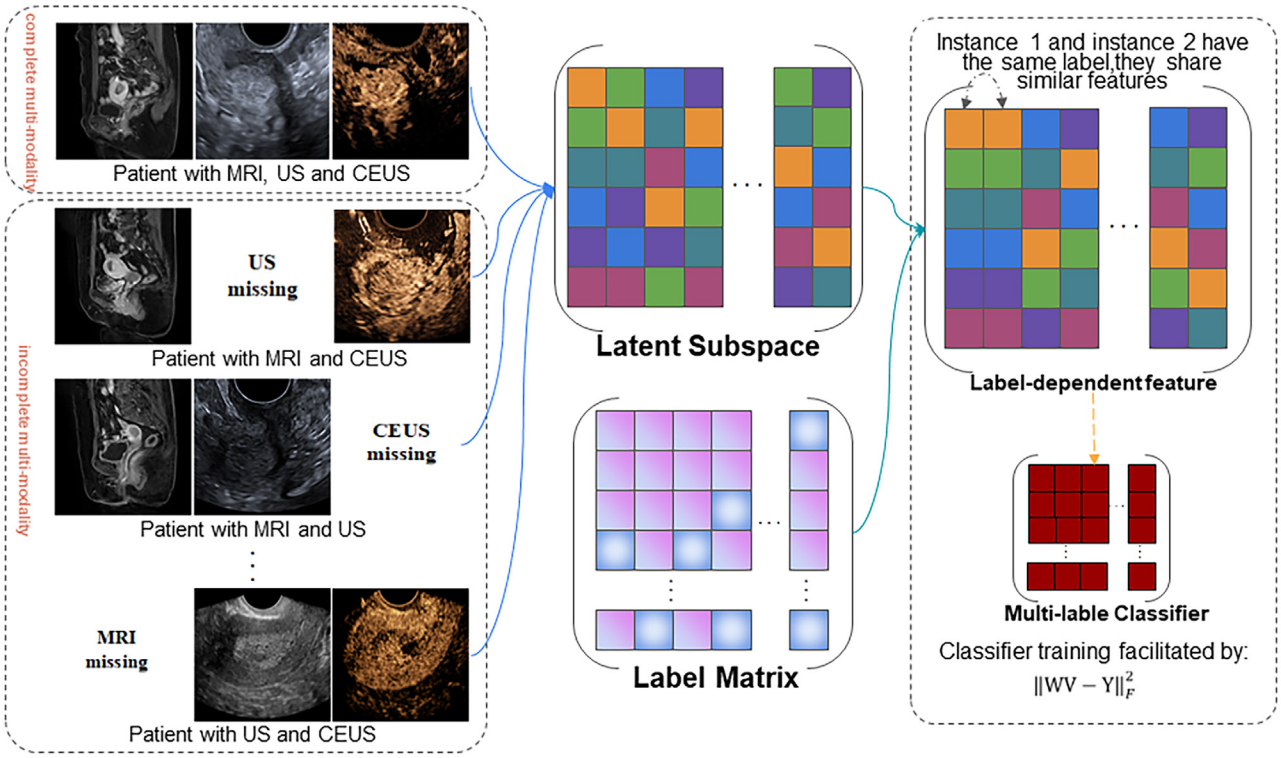


Figure 1. Our proposed latent relation shared frame-work for incomplete multi-modality medical images

To address these challenges, we propose a model that can use incomplete multi-modality medical data, which needs to explore potential relationships between existing modal data, as shown in Figure 1. Specifically, first we map the original modal view into a latent space in order to get different attribute information and reduce the noise of data and the dimensionality of features. Second, latent relation is obtained utilizing samples with incomplete multi-modality images. Finally, mapping all data (complete and incomplete multi-modality) into a potentially unique space. This method shares weights with the potential space, so that it will be able to fill in the information of incomplete modalities. Specifically, input is the incomplete multi-modality medical images, and the output (real label) is endometrial carcinoma stage of that patient. The comparison of our framework with clinical diagnostic method and other methods is shown in Figure 2.

The contributions of this study can be summarized as follows.

(1) A latent relation shared framework for medical images is proposed. It can find the common latent space of multi-view data and perform missing data inference simultaneously to achieve better diagnostic performance.
(2) An incomplete multi-modality approach is proposed, which utilizes image data from all available modalities without introducing extraneous noise parameters, thereby significantly improving prediction performance.
(3) An incomplete multi-modality approach that incorporates MRI, ultrasound and contrast-enhanced ultrasound (CEUS) images is the first attempt at modeling in endometrial cancer detection to the best of our knowledge.
(4) A number of experiments were conducted to evaluate the effectiveness of the proposed method. The proposed model outperforms the current state-of-the-art models on the Tiantan Hospital Endometrial Cancer (TTHEC) task and is suitable for a wider range of disease diagnostic applications.

RESULTS

In this section, we first introduce the experimental settings, including experiment tasks and evaluation strategies. Subsequently, we describe the preprocessing pipeline used in this work. Finally, the performance of different combinations of modalities and incomplete data with different missing rate are showed.

Endometrial carcinoma is divided into four stages according to the International Federation of Gynecology and Obstetrics (FIGO).[3] Compared to stage I and II, stage III and IV are easier to identify on imaging as they are characterized by invading into lower vagina, sidewall, ureters, adjacent, and distant organs. It is more meaningful to distinguish between stage I and II. At the same time, stage IA and IB are subclasses of stage I, and the discrimination of stage IA and IB lesions is also a difficulty in clinical diagnosis. Therefore, this study focus on differentiation of endometrial cancer at stage IA, stage IB and stage II based on incomplete multi-modality medical images, hoping to play an auxiliary role in clinical diagnosis.

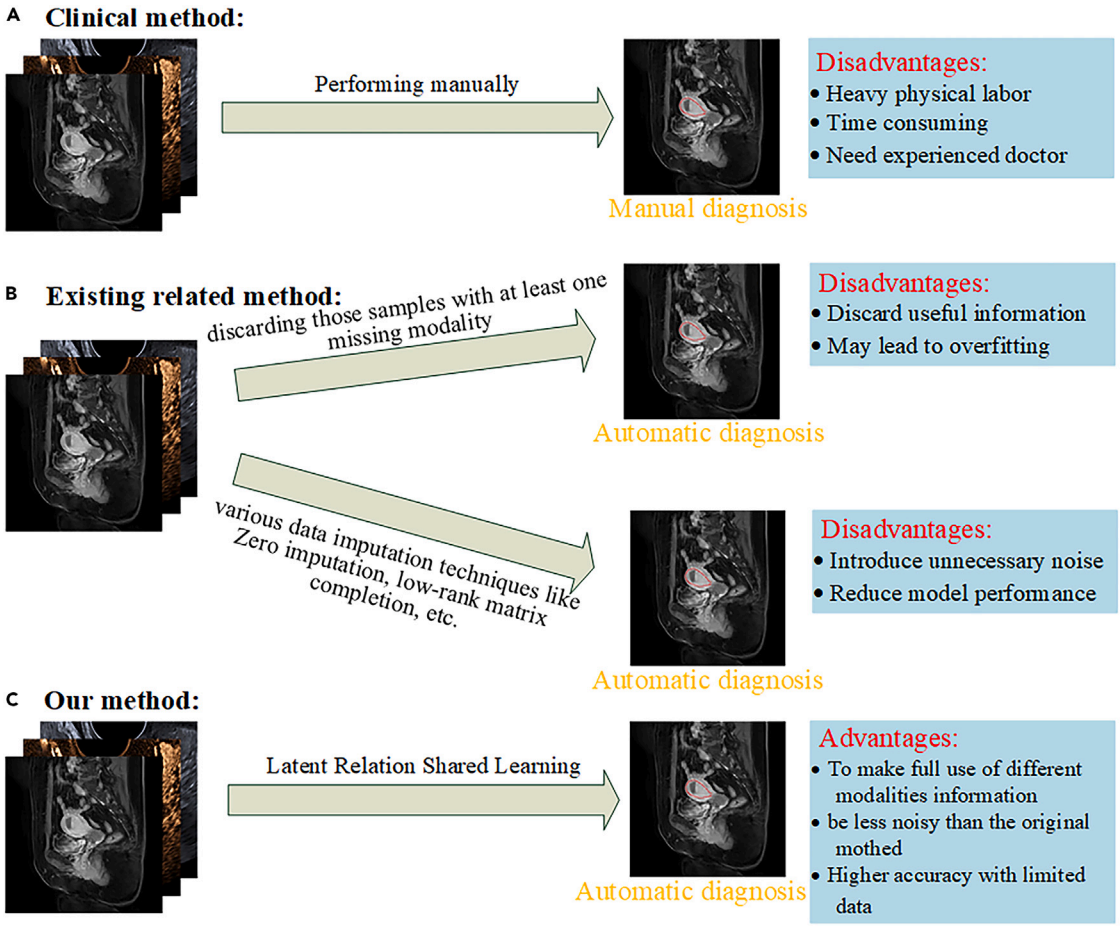


Figure 2. Advantages of our framework

Experimental setup

We apply our method to multi-classification tasks; (1) three-class classification task for normal control (NC) subjects, stage I patients, and stage II patients, (2) four-class classification task for NC subjects, stage IA patients, stage IB patients, and stage II patients, and (3) two-class classification task for stage IA patients and stage IB patients.

To enable a fair comparison, horizontal flipping, vertical flipping, cropping, contrast-modified, and random noise augmentation methods are applied equally to all methods.

Based on the size of the image data, we set up 10-fold cross validation strategy to evaluate all comparison methods. Furthermore, we repeat the experiment for 30 times, and calculate the average value of the results to avoid possible bias. The accuracy (ACC), sensitive (SEN), specificity (SPE), and area under curve (AUC) are used as standard metrics to evaluate the performance of our model.

Related work

Canonical correlation analysis (CCA) is a common dimensionality reduction algorithm,[18] the data are mapped from high dimension to low dimension, and the dispersion of data are ensured as large as possible.[19,20] We extract various features from the image, and each feature can form a linear space. In order to analyze the correlation between these spaces, CCA can be used for analysis.[21] KCCA algorithm improves CCA algorithm by introducing the concept of "kernel trick",[22] realizing the solution of nonlinear problems in the original space and increasing the dimension of the original vector space, and increasing the flexibility.[23,24] The advantages of DCCA over KCCA are that inner product and parameterization methods are not required: the training time varies with the data size, and there is no need to reference the training data when calculating representations where no instances have been seen.[25,26]

Multi-view feature learning requires consensus principle and complementary principle.[27,28] There is paired information between different views of multi-view data. For a multi-view sample, multiple views describe the same object, so there is consistency in multi-view data. The consistency principle is to maximize the consistency of multiple disparate multi-view.[29] In the task of multi-view data, each view should be able to complete a specific knowledge discovery task, and different views usually contain complementary information.[30–32] From an information perspective, in addition to some information shared with other views in each view, there is also some information which is unique to this

view. The existence of complementarity is also the reason why multi-view are superior to single-views. In the early stage of the development of this technology, people often focused on unsupervised dimensionality reduction[22] of multiple views, which showed advantages in clustering and recognition tasks. Recently, probabilistic multi-view graph embedding[33] has been used for multi-view feature learning with many-to-many associations, which is generalizes various existing multi-view methods. Then, the asymmetric merit of LINEX loss[34] is used to improve view-consistency and generalization capability.

Incomplete multi-view feature learning. Recently, a number of incomplete multi-view learning methods[35–40] have been proposed to improve experimental performance. Liu[41] proposed a self-supervised multi-view graph completion algorithm to infer the associated missing entries for incomplete multi-view clustering. An imputation-free deep incomplete multi-view clustering method was proposed, which used an adaptive feature projection to obtain deeper semantic understanding.[42] Xia[43] proposed a kernelized graph-based incomplete multi-view clustering algorithm to overcome the absence of some instances. Incorporating with the constraint of intrinsic geometric structure is proposed to couple incomplete multi-view samples.[44] Data imputation and self-representation learning[45] are jointly conducted to solve the problem of partial data missing.

Taking advantage of the latent relation shared method, we expect to develop an incomplete multi-modality medical images fusion method to supplement the missing information in the mapping space.

Comparison with baseline methods

We compare our method with several baseline methods, including KCCA, DCCA, MDcR, Least_iMSF, Logi_iMSF, and iMVWL. CCA, multi-view feature learning, incomplete multi-view feature learning are selected as the baseline methods. KCCA[46] and DCCA[47] are chosen as the representative methods for CCA. The methods of multi-view feature learning are MDcR.[48] iMSF (Least_iMSF R and Logi_iMSF L)[49] and iMVWL[50] are applied to incomplete multi-view feature learning.

Canonical Correlation Analysis: KCCA maps low-dimensional data to high-dimensional feature space (kernel function space) and performs correlation analysis through kernel function; DCCA is a way of learning complex nonlinear transformations of two views of data whose resulting representations are highly linearly correlated. Multi-view feature learning: MDcR employs Hilbert-Schmidt Independence Criterion to explicitly enforce the dependence across different images. Incomplete multi-view feature learning: iMSF divides the samples into multiple views and constructs a specific classifier for each view. The difference between Least_iMSF and Logi_iMSF is the loss function, which is least squared loss and logical loss in order. iMVWL learns a shared subspace from incomplete views with weak labels, local label correlations, and a predictor in this subspace, simultaneously.

The results of the comparison are shown in Table 1. From the table, we can observe that, in NC vs. stage I vs. stage II, NC vs. stage IA vs. stage IB vs. stage II and stage IA vs. stage IB classification, "OURS" generally outperforms the other methods in terms of ACC, SEN, SPE, and AUC, which implies that more feature learned by our proposed method can help improve the classification performance. Second, the results reported for multi-classification tasks (NC vs. stage I vs. stage II, NC vs. stage IA vs. stage IB vs. stage II) are lower than the performance on the binary classification task (stage IA vs. stage IB). The most likely reason is that multi-classification tasks are more challenging than binary classification tasks in medical image analysis. There are some similarities in the morphology of the different stages. Besides, using all available data methods ("Ours", "Least_iMSF", "Logi_iMSF", and "iMVWL") is better than using complete multi-modality data methods ("KCCA", "DCCA", "MDcR"). The results of OURS (complete multi-modality data) are superior or equal to that of Least_iMSF and Logi_iMSF, but inferior to iMVWL. Our model has certain advantages in data feature extraction, but the discarded data will affect the performance of the model. In addition, the incomplete image information contains useful medical diagnostic data, which can be used to improve the performance of the model.

Study on different combinations of modalities

This part mainly studies the latent features of different modalities and the correlations between them. We compare the performance of our method for different combinations of modalities on different classification tasks in Table 2. As can be seen from the Table 2, our model using MRI outperforms the models using other two kind of modalities (US, CEUS), and compared with the data of the other two modalities, the discrimination ability of US data in patients with endometrial cancer may be poor. Similarly, MRI+CEUS performances better than MRI+US. This implies that MRI could be comparatively more effective biomarkers in distinguishing endometrial cancer status diagnosis compared to US and CEUS data. However, MRI, US, and CEUS still contain their specific information, and we can observe that the "MRI+US+CEUS" modal outperforms other combinations for the NC vs. stage I vs. stage II, NC vs. stage IA vs. stage IB vs. stage II, and stage IA vs. stage IB classification tasks.

Study on incomplete data with different missing rates

MRI, US, and CEUS data are not available for all patients in the diagnosis of endometrial disease. In order to further verify the effectiveness of the proposed method in dealing with different missing modality, partial data of MRI, US, or CEUS were randomly discarded, with 10%, 20%, and 30% of each data discarded, respectively. The performance is showed in Table 3.

As the missing rate increases from 10% to 30%, it was found that the accuracy of all tasks decreases as the miss rate increased, although the rate of decline was different for different methods. As shown in Table 3, our proposed method outperforms almost all the other methods under different missing rates.

Overall, the resolution of endometrial cancer stage IA vs. stage IB is higher than that of NC vs. stage I vs. stage II and NC vs. stage IA vs. stage IB vs. stage II tasks at different missing rates. As NC and stage I are similar in imaging morphology, so the performance of the

Table 1. Classification performance of all compared methods

Tasks	Method	ACC(%)	SEN(%)	SPE(%)	AUC
NC vs. stage I vs. stage II	KCCA	55.2 ± 0.85	51.3 ± 1.85	64.3 ± 0.97	0.577 ± 0.011
	DCCA	57.1 ± 0.73	52.7 ± 2.31	65.1 ± 0.85	0.590 ± 0.024
	MDcR	61.1 ± 0.62	59.2 ± 2.12	68.5 ± 0.82	0.673 ± 0.009
	Least_iMSF	72.5 ± 0.92	60.1 ± 2.89	88.3 ± 0.88	0.719 ± 0.009
	Logi_iMSF	72.0 ± 0.86	59.8 ± 2.94	88.1 ± 0.76	0.702 ± 0.010
	iMVWL	79.8 ± 0.65	76.8 ± 2.77	87.5 ± 0.89	0.813 ± 0.006
	OURS(complete multi-modality data)	73.1 ± 0.77	60.0 ± 2.55	88.1 ± 0.83	0.710 ± 0.008
	OURS	80.1 ± 0.81	79.6 ± 2.67	89.2 ± 0.76	0.823 ± 0.008
NC vs. stage IA vs. stage IB vs. stage II	KCCA	51.7 ± 0.81	48.8 ± 1.93	63.1 ± 1.02	0.523 ± 0.015
	DCCA	53.3 ± 0.79	50.1 ± 2.75	63.2 ± 0.93	0.541 ± 0.030
	MDcR	55.9 ± 0.68	51.0 ± 2.33	62.5 ± 0.95	0.582 ± 0.018
	Least_iMSF	68.1 ± 0.87	54.8 ± 2.95	86.8 ± 0.98	0.680 ± 0.018
	Logi_iMSF	67.3 ± 0.95	54.9 ± 2.99	86.2 ± 0.95	0.669 ± 0.017
	iMVWL	76.9 ± 0.52	75.3 ± 2.89	85.3 ± 1.03	0.785 ± 0.023
	OURS(complete multi-modality data)	68.1 ± 0.72	59.8 ± 2.81	86.6 ± 0.92	0.706 ± 0.010
	OURS	78.2 ± 0.79	75.3 ± 2.82	87.1 ± 0.85	0.791 ± 0.009
stage IA vs. stage IB	KCCA	63.5 ± 0.88	59.8 ± 1.62	75.6 ± 0.93	0.661 ± 0.007
	DCCA	64.1 ± 0.76	61.2 ± 2.42	76.1 ± 0.81	0.670 ± 0.018
	MDcR	68.3 ± 0.61	65.3 ± 2.25	77.0 ± 0.82	0.741 ± 0.008
	Least_iMSF	72.8 ± 0.91	60.5 ± 2.75	89.1 ± 0.85	0.722 ± 0.006
	Logi_iMSF	71.5 ± 0.95	60.1 ± 2.91	89.0 ± 0.72	0.719 ± 0.007
	iMVWL	82.5 ± 0.63	79.2 ± 2.85	88.9 ± 0.69	0.828 ± 0.007
	OURS(complete multi-modality data)	75.2 ± 0.70	78.3 ± 2.52	89.2 ± 0.81	0.753 ± 0.008
	OURS	84.1 ± 0.72	82.9 ± 2.73	90.1 ± 0.76	0.877 ± 0.006

Results are shown as mean ± standard deviation.

model will be affected. At the same time, the multi-detection problem of NC vs. stage IA vs. stage IB vs. stage II reduces the efficiency of the model. At the same missing rate, MRI has the greatest impact on the model, followed by CEUS and US. It shows that in the process of model building, the MRI maps the most effective information to latent space among the three kind of modal medical images. Discarding partial data in any medical image will lead to a decrease in the performance of the model, indicating that all three medical images map useful information to the latent space, and the information is complementary. This part of the content is confirmed by the

Table 2. Classification results of different modality combinations for three classification tasks

	NC vs. stage I vs. stage II		NC vs. stage IA vs. stage IB vs. stage II		stage IA vs. stage IB	
Modalities	ACC(%)	AUC	ACC(%)	AUC	ACC(%)	AUC
MRI	72.1 ± 1.02	0.818 ± 0.004	69.5 ± 1.49	0.798 ± 0.009	75.8 ± 0.98	0.862 ± 0.009
US	70.1 ± 1.96	0.805 ± 0.006	63.3 ± 2.21	0.812 ± 0.010	74.7 ± 1.95	0.812 ± 0.012
CEUS	71.3 ± 1.83	0.815 ± 0.013	67.7 ± 1.99	0.799 ± 0.007	74.8 ± 1.90	0.829 ± 0.006
MRI+US	78.5 ± 1.10	0.818 ± 0.006	77.2 ± 1.32	0.803 ± 0.008	83.1 ± 1.01	0.862 ± 0.008
MRI+CEUS	78.9 ± 0.93	0.817 ± 0.010	77.5 ± 1.02	0.790 ± 0.006	83.5 ± 0.89	0.870 ± 0.006
MRI+US+CEUS	80.1 ± 0.81	0.823 ± 0.008	78.2 ± 0.79	0.791 ± 0.009	84.1 ± 0.72	0.877 ± 0.006

Results are shown as mean ± standard deviation.

Table 3. Classification performance of all compared methods under different missing rates

Method	Missing rate	NC vs. stage I vs. stage II		NC vs. stage IA vs. stage IB vs. stage II		stage IA vs. stage IB	
		ACC(%)	AUC	ACC(%)	AUC	ACC(%)	AUC
(a) Missing MRI							
Least_iMSF	10%	69.8 ± 1.03	0.720 ± 0.010	66.0 ± 1.58	0.685 ± 0.013	70.9 ± 1.12	0.725 ± 0.007
	20%	69.1 ± 1.09	0.716 ± 0.013	64.9 ± 1.89	0.702 ± 0.020	69.3 ± 1.33	0.730 ± 0.015
	30%	67.3 ± 1.52	0.680 ± 0.007	64.1 ± 2.10	0.694 ± 0.019	67.5 ± 1.57	0.700 ± 0.008
Logi_iMSF	10%	70.1 ± 0.96	0.710 ± 0.009	65.8 ± 1.21	0.628 ± 0.032	70.2 ± 1.18	0.683 ± 0.013
	20%	69.1 ± 1.22	0.712 ± 0.023	64.5 ± 1.98	0.620 ± 0.028	68.4 ± 1.72	0.681 ± 0.011
	30%	67.5 ± 1.49	0.701 ± 0.017	62.0 ± 2.35	0.600 ± 0.019	67.3 ± 1.85	0.675 ± 0.007
iMVWL	10%	77.8 ± 0.72	0.812 ± 0.011	73.5 ± 0.80	0.778 ± 0.012	80.5 ± 0.79	0.825 ± 0.019
	20%	76.5 ± 0.92	0.810 ± 0.009	71.5 ± 1.21	0.790 ± 0.015	78.2 ± 0.93	0.811 ± 0.022
	30%	75.2 ± 1.13	0.799 ± 0.015	69.3 ± 1.62	0.761 ± 0.009	75.8 ± 1.21	0.781 ± 0.013
OURS	10%	78.1 ± 0.92	0.825 ± 0.007	74.5 ± 1.12	0.781 ± 0.013	82.0 ± 0.90	0.879 ± 0.008
	20%	77.3 ± 1.12	0.813 ± 0.009	72.3 ± 1.18	0.796 ± 0.016	79.8 ± 1.02	0.816 ± 0.006
	30%	75.9 ± 1.18	0.799 ± 0.005	71.5 ± 1.39	0.775 ± 0.017	78.0 ± 1.20	0.800 ± 0.018
(b) Missing US							
Least_iMSF	10%	71.6 ± 0.82	0.716 ± 0.011	67.5 ± 1.20	0.673 ± 0.015	72.2 ± 0.91	0.716 ± 0.006
	20%	71.0 ± 0.97	0.715 ± 0.015	66.7 ± 1.25	0.663 ± 0.017	71.5 ± 1.19	0.700 ± 0.019
	30%	70.1 ± 1.01	0.680 ± 0.009	66.1 ± 1.55	0.669 ± 0.016	70.5 ± 1.24	0.698 ± 0.008
Logi_iMSF	10%	71.7 ± 0.85	0.701 ± 0.007	67.1 ± 1.28	0.645 ± 0.013	71.2 ± 0.82	0.710 ± 0.019
	20%	70.9 ± 1.07	0.700 ± 0.026	66.4 ± 1.35	0.638 ± 0.015	70.9 ± 1.00	0.689 ± 0.017
	30%	70.2 ± 1.09	0.675 ± 0.016	66.1 ± 1.68	0.635 ± 0.009	70.6 ± 1.01	0.679 ± 0.012
iMVWL	10%	79.3 ± 0.65	0.812 ± 0.015	76.0 ± 0.76	0.782 ± 0.010	81.6 ± 0.70	0.820 ± 0.011
	20%	78.5 ± 0.83	0.809 ± 0.019	74.8 ± 0.96	0.778 ± 0.011	80.7 ± 0.78	0.815 ± 0.008
	30%	77.6 ± 0.97	0.779 ± 0.017	74.1 ± 1.32	0.769 ± 0.012	79.8 ± 0.96	0.789 ± 0.016
OURS	10%	79.5 ± 0.85	0.820 ± 0.013	77.3 ± 0.89	0.785 ± 0.010	82.9 ± 0.86	0.870 ± 0.009
	20%	78.9 ± 0.88	0.816 ± 0.015	76.1 ± 1.02	0.773 ± 0.009	81.7 ± 0.96	0.846 ± 0.005
	30%	78.0 ± 0.98	0.814 ± 0.005	75.3 ± 1.10	0.773 ± 0.017	80.3 ± 1.02	0.828 ± 0.008
(c) Missing CEUS							
Least_iMSF	10%	71.0 ± 0.85	0.715 ± 0.015	67.3 ± 1.33	0.665 ± 0.011	71.2 ± 1.03	0.710 ± 0.013
	20%	70.1 ± 1.01	0.708 ± 0.013	66.0 ± 1.49	0.659 ± 0.008	70.3 ± 1.33	0679 ± 0.010
	30%	68.9 ± 1.25	0.682 ± 0.019	64.5 ± 1.92	0.635 ± 0.015	68.9 ± 1.45	0.672 ± 0.018
Logi_iMSF	10%	71.0 ± 0.95	0.698 ± 0.009	66.0 ± 1.25	0.654 ± 0.013	70.4 ± 1.09	0.700 ± 0.015
	20%	69.9 ± 1.10	0.689 ± 0.011	65.1 ± 1.33	0.630 ± 0.012	68.9 ± 1.35	0.685 ± 0.008
	30%	69.0 ± 1.35	0.663 ± 0.012	62.9 ± 1.91	0.619 ± 0.010	68.0 ± 1.32	0.671 ± 0.012
iMVWL	10%	79.3 ± 0.70	0.802 ± 0.011	74.2 ± 0.76	0.779 ± 0.012	81.1 ± 0.69	0.823 ± 0.015
	20%	78.2 ± 0.90	0.793 ± 0.015	71.9 ± 0.92	0.770 ± 0.013	78.5 ± 0.75	0.813 ± 0.012
	30%	77.3 ± 0.92	0.776 ± 0.013	69.9 ± 0.99	0.751 ± 0.007	76.6 ± 1.01	0.785 ± 0.007
OURS	10%	79.3 ± 0.86	0.816 ± 0.009	75.3 ± 1.03	0.783 ± 0.011	82.3 ± 0.90	0.869 ± 0.009
	20%	78.5 ± 0.95	0.810 ± 0.006	72.5 ± 1.32	0.768 ± 0.009	80.1 ± 1.17	0.833 ± 0.013
	30%	77.8 ± 0.99	0.801 ± 0.011	72.0 ± 1.40	0.770 ± 0.015	78.3 ± 1.32	0.810 ± 0.015

Results are shown as mean ± standard deviation.

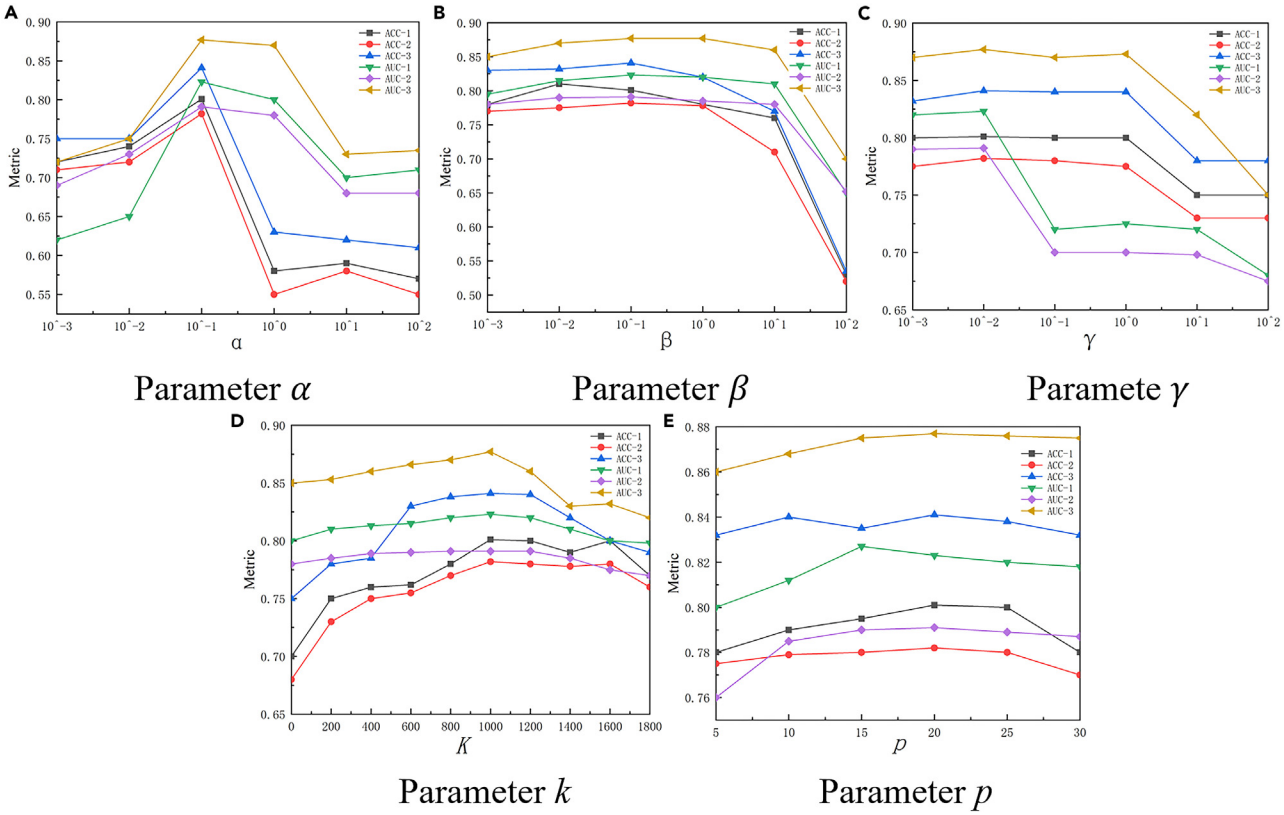


Figure 3. Performance of each parameter

conclusion of different combinations of modalities, indicating that the conclusion is still applicable under the conditions of different missing rates.

As can be seen from Table 3, our model has certain advantages under the conditions of three different images missing. Comparing stage IA vs. stage IB and NC vs. stage IA vs. stage IB vs. stage II, our model has a certain decline in the performance of two-class and four-class detection, but it still has advantages compared to other models. The possible reason may be that limited patient data are used as much as possible in model training. Although the accuracy of the model decreases with the increase of missing rate, our model still achieves a high level of accuracy in three different classification tasks despite the missing rate reached 30%.

Ablation study and parameters analysis

In order to investigate the effect of parameters and demonstrate the effectiveness of the Laplacian regularization term for our approach for endometrial cancer diagnosis with incomplete multi-modality medical images, we analyzed the five employed hyperparameters α, β, γ, k and p, respectively, and conducted model simplification tests.

Firstly, the effect of trade-off parameters, dimension of the latent subspace and neighbor number on the model are studied, the results are shown in Figure 3. ACC-1, ACC-2, and ACC-3 represent the accuracy of tasks NC vs. stage I vs. stage II, NC vs. stage IA vs. stage IB vs. stage II, and stage IA vs. stage IB, respectively. AUC-1, AUC-2, and AUC-3 represent the under curve of tasks NC vs. stage I vs. stage II, NC vs. stage IA vs. stage IB vs. stage II, and stage IA vs. stage IB, respectively. α, β, and γ are the trade-off parameters of our model. In optimization problems, trade-off parameters are used to balance different objectives. We set the parameters α, β, and γ to 10^-1, 10^-1, and 10^-2, respectively. Subspace dimension k contains important latent subspace information, which we set to 1000. P is the neighbor number. From the performance in Figure 3, we set p to 20.

Secondly, to verify the effectiveness of the Laplacian regularization term, we conducted the ablation study between "OURS-L" and "OURS" to remove the Laplacian regularization process from the model. Table 4 shows the comparative results of the four metrics. It can be found that when the Laplacian regularization process is removed, the performance of the model decreases significantly. And some of the results are even worse than the method using complete multi-modality data. Experimental results show that Laplacian regularization helps to maintain smoothness of data in learning and accelerate the model's understanding and representation learning of graph structure. At the same time, it can reduce overfitting and improve generalization.

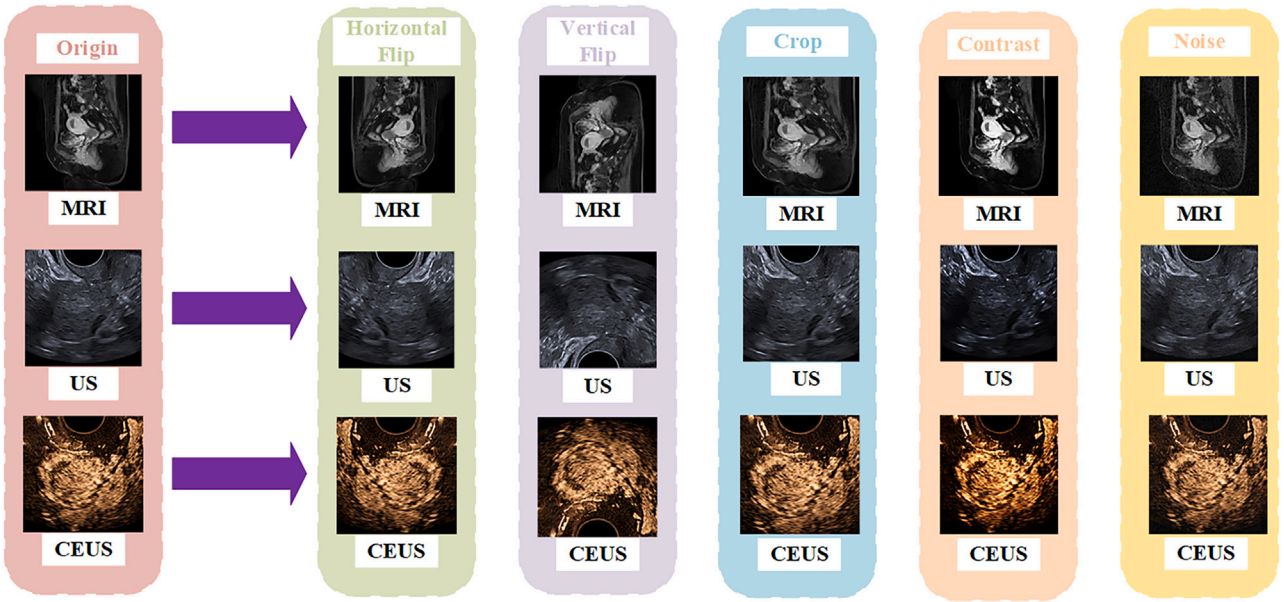


Figure 4. Dataset Augmentation

DISCUSSION

In this paper, we propose a framework to identify benign and early-stage endometrial cancer. To this end, we propose a latent relation shared representation learning method for incomplete multi-modality medical images fusion. The idea is to map all complete modal samples and incomplete modal samples into the potential mapping space, and use this method to "share" information, so that potential connections can be explored. Moreover, the proposed method makes full use of clinical incomplete multi-modality images data to assist clinical diagnosis and staging.

Our framework achieves accuracy of 80.1%, 78.2%, and 84.1% on target class of NC vs. stage I vs. stage II, NC vs. stage IA vs. stage IB vs. stage II and stage IA vs. stage IB, which outperforms all six baseline methods. Meanwhile, we find "MRI+US+CEUS" modal performs best among all combinations. Furthermore, the performance of our method is also excellent under different imaging missing rates. In the most difficult classification task NC vs. stage IA vs. stage IB vs. stage II, even when the missing rate is 30%, the accuracy of our model on missing MRI, missing US, and missing CEUS can reach 0.715, 0.753, and 0.720, respectively.

Limitations of the study

However, there are some weak points of our methods. Only information related to endometrial cancer was annotated in the dataset, but no other diseases were considered. In the future work, we hope to design a fusion method that considers the correlation between endometrial

Table 4. Results of OURS-L and OURS on the three tasks

Tasks	Metrics	OURS-L	OURS
NC vs. stage I vs. stage II	ACC(%)	58.3 ± 0.89	80.1 ± 0.81
	SEN(%)	53.6 ± 2.53	79.6 ± 2.67
	SPE(%)	69.3 ± 0.92	89.2 ± 0.76
	AUC	0.617 ± 0.018	0.823 ± 0.008
NC vs. stage IA vs. stage IB vs. stage II	ACC(%)	51.2 ± 0.75	78.2 ± 0.79
	SEN(%)	52.6 ± 2.77	75.3 ± 2.82
	SPE(%)	65.2 ± 1.02	87.1 ± 0.85
	AUC	0.545 ± 0.014	0.791 ± 0.009
stage IA vs. stage IB	ACC(%)	67.7 ± 0.42	84.1 ± 0.72
	SEN(%)	68.1 ± 2.13	82.9 ± 2.73
	SPE(%)	80.3 ± 0.91	90.1 ± 0.76
	AUC	0.723 ± 0.013	0.877 ± 0.006

Results are shown as mean ± standard deviation.

Table 5. Detailed information of datasets

	stage IA	stage IB	stage II
Size(Number of patients)	43	36	38
Missing rate(MRI)	2.3%	0	0
Missing rate(US)	0	2.8%	0
Missing rate(CEUS)	7.0%	5.6%	5.3%

stage IA and IB are subclasses of stage I.

cancer and other diseases to further improve the performance of endometrial cancer classification. Moreover, we also plan to fuse medical image data with medical text data to improve the accuracy of multi-classification problems in the case of incomplete multi-modality.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Dataset Augmentation
  - Mapping for multi-modality representation learning
  - Incomplete multi-modality representation learning
  - Optimization
- QUANTIFICATION AND STATISTICAL ANALYSIS

ACKNOWLEDGMENTS

This work is supported by the National Key R&D Program of China (No. 2020AAA0106600).

AUTHOR CONTRIBUTIONS

J.L.: methodology, writing – original draft. L.L.: supervision, reviewing and editing. M.J.: reviewing and editing. Z.C.: reviewing and editing. X.L.: supervision, reviewing and editing.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 25, 2024
Revised: May 22, 2024
Accepted: July 11, 2024
Published: July 15, 2024

REFERENCES

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. (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. https://doi.org/10.3322/caac.21492.
2. Asami, Y., Kobayashi Kato, M., Hiranuma, K., Matsuda, M., Shimada, Y., Ishikawa, M., Koyama, T., Komatsu, M., Hamamoto, R., Nagashima, M., et al. (2023). Utility of molecular subtypes and genetic alterations for evaluating clinical outcomes in 1029 patients with endometrial cancer. Br. J. Cancer *128*, 1582–1591. https://doi.org/10.1038/s41416-023-02203-3.
3. Amant, F., Mirza, M.R., Koskas, M., and Creutzberg, C.L. (2018). Cancer of the corpus uteri. Int. J. Gynaecol. Obstet. *143*, 37–50. https://doi.org/10.1002/ijgo.12612.
4. Cade, T.J., Quinn, M.A., McNally, O.M., Neesham, D., Pyman, J., and Dobrotwir, A. (2010). Predictive value of magnetic resonance imaging in assessing myometrial invasion in endometrial cancer: is radiological staging sufficient for planning conservative treatment? Int. J. Gynecol. Cancer *20*, 1166–1169. https://doi.org/10.1111/igc.0b013e3181e9509f.
5. Ozturk, E., Dikensoy, E., Balat, O., Ugur, M.G., and Aydin, A. (2012). Intraoperative frozen section is essential for assessment of myometrial invasion but not for histologic grade confirmation in endometrial cancer: a ten-year experience. Arch. Gynecol. Obstet. *285*, 1415–1419. https://doi.org/10.1007/s00404-011-2135-z.
6. Karlsson, B., Granberg, S., Wikland, M., Ylostalo, P., Torvid, K., Marsal, K., and Valentin, L. (1995). Transvaginal ultrasonography of the endometrium in

women with postmenopausal bleeding–a Nordic multicenter study. Am. J. Obstet. Gynecol. *172*, 1488–1494. https://doi.org/10.1016/0002-9378(95)90483-2.

7. Zhou, H., He, L., Zhang, Y., Shen, L., and Chen, B. (2022). Interpretable Graph Convolutional Network of Multi-Modality Brain Imaging for Alzheimer's Disease Diagnosis (IEEE), pp. 1–5.
8. Xing, X., Chen, Z., Hou, Y., and Yuan, Y. (2023). Gradient modulated contrastive distillation of low-rank multi-modal knowledge for disease diagnosis. Med. Image Anal. *88*, 102874. https://doi.org/10.1016/j.media.2023.102874.
9. Yu, B., Chen, H., Jia, C., Zhou, H., Cong, L., Li, X., Zhuang, J., and Cong, X. (2023). Multi-modality multi-scale cardiovascular disease subtypes classification using Raman image and medical history. Expert Syst. Appl. *224*, 119965. https://doi.org/10.1016/j.eswa.2023.119965.
10. Cao, Z., Sun, C., Wang, W., Zheng, X., Wu, J., and Gao, H. (2021). Multi-modality fusion learning for the automatic diagnosis of optic neuropathy. Pattern Recogn. Lett. *142*, 58–64.
11. Zhou, J., Zhang, X., Zhu, Z., Lan, X., Fu, L., Wang, H., and Wen, H. (2022). Cohesive multi-modality feature learning and fusion for COVID-19 patient severity prediction. IEEE Trans. Circuits Syst. Video Technol. *32*, 2535–2549.
12. Amini, M., Nazari, M., Shiri, I., Hajianfar, G., Deevband, M.R., Abdollahi, H., Arabi, H., Rahmim, A., and Zaidi, H. (2021). Multi-level multi-modality (PET and CT) fusion radiomics: prognostic modeling for non-small cell lung carcinoma. Phys. Med. Biol. *66*, 205017.
13. Diao, Y., Li, F., and Li, Z. (2023). Joint learning-based feature reconstruction and enhanced network for incomplete multi-modal brain tumor segmentation. Comput. Biol. Med. *163*, 107234. https://doi.org/10.1016/j.compbiomed.2023.107234.
14. Zhu, Q., Li, H., Ye, H., Zhang, Z., Wang, R., Fan, Z., and Zhang, D. (2022). Incomplete multi-modal brain image fusion for epilepsy classification. Inf. Sci. *582*, 316–333. https://doi.org/10.1016/j.ins.2021.09.035.
15. Ren, C.X., Xu, G.X., Dai, D.Q., Lin, L., Sun, Y., and Liu, Q.S. (2024). Cross-site prognosis prediction for nasopharyngeal carcinoma from incomplete multi-modal data. Med. Image Anal. *93*, 103103. https://doi.org/10.1016/j.media.2024.103103.
16. Zhang, D., Wang, C., Chen, T., Chen, W., and Shen, Y. (2024). Scalable Swin Transformer network for brain tumor segmentation from incomplete MRI modalities. Artif. Intell. Med. *149*, 102788. https://doi.org/10.1016/j.artmed.2024.102788.
17. Cao, R., Mohammadian Bajgiran, A., Afshari Mirak, S., Shakeri, S., Zhong, X., Enzmann, D., Raman, S., and Sung, K. (2019). Joint Prostate Cancer Detection and Gleason Score Prediction in mp-MRI via FocalNet. IEEE Trans. Med. Imaging *38*, 2496–2506. https://doi.org/10.1109/TMI.2019.2901928.
18. Hotelling, H. (1992). Relations between two sets of variates. In Breakthroughs in statistics (Springer), pp. 162–190.
19. González, I., Déjean, S., Martin, P.G., and Baccini, A. (2008). CCA: An R package to extend canonical correlation analysis. J. Stat. Software *23*, 1–14.
20. Dhillon, P., Foster, D.P., and Ungar, L. (2011). Multi-view learning of word embeddings via cca. Adv. Neural Inf. Process. Syst. *24*.
21. Yang, X., Liu, W., Liu, W., and Tao, D. (2021). A survey on canonical correlation analysis. IEEE Trans. Knowl. Data Eng. *33*, 2349–2368.
22. Zheng, W., Zhou, X., Zou, C., and Zhao, L. (2006). Facial expression recognition using kernel canonical correlation analysis (KCCA). IEEE Trans. Neural Netw. *17*, 233–238.
23. Bai, Y., Tang, P., and Hu, C. (2018). kCCA transformation-based radiometric normalization of multi-temporal satellite images. Rem. Sens. *10*, 432.
24. Chen, Q., and Wang, Y. (2021). Key-performance-indicator-related state monitoring based on kernel canonical correlation analysis. Control Eng. Pract. *107*, 104692. https://doi.org/10.1016/j.conengprac.2020.104692.
25. Vieluf, S., Hasija, T., Kuschel, M., Reinsberger, C., and Loddenkemper, T. (2023). Developing a deep canonical correlation-based technique for seizure prediction. Expert Syst. Appl. *234*, 120986. https://doi.org/10.1016/j.eswa.2023.120986.
26. Prabhakar, G.A., Basel, B., Dutta, A., and Rao, C.V.R. (2023). Multichannel CNN-BLSTM Architecture for Speech Emotion Recognition System by Fusion of Magnitude and Phase Spectral Features Using DCCA for Consumer Applications. IEEE Trans. Consumer Electron. *69*, 226–235. https://doi.org/10.1109/TCE.2023.3236972.
27. Fang, Y., Xu, F., Wei, L., Jiang, Y., Chen, J., Wei, L., and Wei, D.-Q. (2023). AFP-MFL: accurate identification of antifungal peptides using multi-view feature learning. Brief. Bioinform. *24*, bbac606. https://doi.org/10.1093/bib/bbac606.
28. Chen, Z., Fu, L., Yao, J., Guo, W., Plant, C., and Wang, S. (2023). Learnable graph convolutional network and feature fusion for multi-view learning. Inf. Fusion *95*, 109–119. https://doi.org/10.1016/j.inffus.2023.02.013.
29. Li, J., Zhang, B., Lu, G., and Zhang, D. (2019). Generative multi-view and multi-feature learning for classification. Inf. Fusion *45*, 215–226.
30. Tian, X., Deng, Z., Ying, W., Choi, K.-S., Wu, D., Qin, B., Wang, J., Shen, H., and Wang, S. (2019). Deep multi-view feature learning for EEG-based epileptic seizure detection. IEEE Trans. Neural Syst. Rehabil. Eng. *27*, 1962–1972.
31. Wang, H., Nie, F., and Huang, H. (2013). Multi-view Clustering and Feature Learning via Structured Sparsity (PMLR), pp. 352–360.
32. Xue, X., Nie, F., Wang, S., Chang, X., Stantic, B., and Yao, M. (2017). Multi-view Correlated Feature Learning by Uncovering Shared Component, p. 1.
33. Okuno, A., Hada, T., and Shimodaira, H. (2018). A Probabilistic Framework for Multi-View Feature Learning with Many-To-Many Associations via Neural Networks (PMLR), pp. 3888–3897.
34. Tang, J., Xu, W., Li, J., Tian, Y., and Xu, S. (2021). Multi-view learning methods with the LINEX loss for pattern classification. Knowl. Base Syst. *228*, 107285.
35. Zhang, W., Deng, Z., Zhang, T., Choi, K.-S., Wang, J., and Wang, S. (2021). Incomplete multi-view fuzzy inference system with missing view imputation and cooperative learning. IEEE Trans. Fuzzy Syst. *30*, 3038–3051.
36. Wen, J., Zhang, Z., Zhang, Z., Zhu, L., Fei, L., Zhang, B., and Xu, Y. (2021). Unified tensor framework for incomplete multi-view clustering and missing-view inferring. Proc. Int. AAAI Conf. Weblogs Soc. Media *35*, 10273–10281.
37. Zhu, P., Yao, X., Wang, Y., Cao, M., Hui, B., Zhao, S., and Hu, Q. (2022). Latent Heterogeneous Graph Network for Incomplete Multi-View Learning. In IEEE Transactions on Multimedia.
38. Wen, J., Liu, C., Deng, S., Liu, Y., Fei, L., Yan, K., and Xu, Y. (2023). Deep Double Incomplete Multi-View Multi-Label Learning With Incomplete Labels and Missing Views. In IEEE Transactions on Neural Networks and Learning Systems, pp. 1–13. https://doi.org/10.1109/TNNLS.2023.3260349.
39. Deng, S., Wen, J., Liu, C., Yan, K., Xu, G., and Xu, Y. (2023). Projective Incomplete Multi-View Clustering. In IEEE Transactions on Neural Networks and Learning Systems, pp. 1–13. https://doi.org/10.1109/TNNLS.2023.3242473.
40. Zhou, T., Liu, M., Thung, K.H., and Shen, D. (2019). Latent Representation Learning for Alzheimer's Disease Diagnosis With Incomplete Multi-Modality Neuroimaging and Genetic Data. IEEE Trans. Med. Imaging *38*, 2411–2422. https://doi.org/10.1109/TMI.2019.2913158.
41. Liu, C., Wu, S., Li, R., Jiang, D., and Wong, H.S. (2023). Self-Supervised Graph Completion for Incomplete Multi-View Clustering. IEEE Trans. Knowl. Data Eng. *35*, 9394–9406. https://doi.org/10.1109/TKDE.2023.3238416.
42. Xu, J., Li, C., Peng, L., Ren, Y., Shi, X., Shen, H.T., and Zhu, X. (2023). Adaptive Feature Projection With Distribution Alignment for Deep Incomplete Multi-View Clustering. IEEE Trans. Image Process. *32*, 1354–1366. https://doi.org/10.1109/TIP.2023.3243521.
43. Xia, D., Yang, Y., Yang, S., and Li, T. (2023). Incomplete multi-view clustering via kernelized graph learning. Inf. Sci. *625*, 1–19. https://doi.org/10.1016/j.ins.2023.01.013.
44. Zhao, L., Chen, Z., Yang, Y., Wang, Z.J., and Leung, V.C. (2018). Incomplete multi-view clustering via deep semantic mapping. Neurocomputing *275*, 1053–1062.
45. Liu, J., Liu, X., Zhang, Y., Zhang, P., Tu, W., Wang, S., Zhou, S., Liang, W., Wang, S., and Yang, Y. (2021). Self-representation Subspace Clustering for Incomplete Multi-View Data, pp. 2726–2734.
46. Akaho, S. (2006). A Kernel Method for Canonical Correlation Analysis. Preprint at arXiv. https://doi.org/10.48550/arXiv.cs/0609071.
47. Andrew, G., Arora, R., Bilmes, J., and Livescu, K. (2013). Deep Canonical Correlation Analysis (PMLR), pp. 1247–1255.
48. Zhang, C., Fu, H., Hu, Q., Zhu, P., and Cao, X. (2017). Flexible Multi-View Dimensionality Co-Reduction. IEEE Trans. Image Process. *26*, 648–659. https://doi.org/10.1109/TIP.2016.2627806.
49. Yuan, L., Wang, Y., Thompson, P.M., Narayan, V.A., and Ye, J.; Alzheimer's Disease Neuroimaging (2012). Multi-source feature learning for joint analysis of incomplete multiple heterogeneous neuroimaging data. Neuroimage *61*, 622–632. https://doi.org/10.1016/j.neuroimage.2012.03.059.
50. Tan, Q., Yu, G., Domeniconi, C., Wang, J., and Zhang, Z. (2018). Incomplete Multi-View Weak-Label Learning, pp. 2703–2709.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Research Data Deposit	Beijing Tiantan Hospital	Available upon request following the completion of a suitable confidentiality agreement. (No.TTHEC-KYSQ 202212401)
Software and algorithms		
Python (version 3.7.0)	Python software	https://www.python.org/
CUDA	Version 9.2	https://developer.nvidia.com/cuda-92-download-archive
Matplotlib	Version 3.3.1	https://matplotlib.org/3.3.1/
Prediction models	This study	https://github.com/jq-gogogo/Latent_Relation_Shared_Learning
KCCA	Akaho et al.[46]	https://github.com/lorenzoriano/PyKCCA
DCCA	Andrew et al.[47]	https://github.com/Michaelvll/DeepCCA
MDcR	Zhang et al.[48]	http://cic.tju.edu.cn/faculty/zhangchangqing/code.html
Least_iMSF	Yuan et al.[49]	https://github.com/jiayuzhou/MALSAR/tree/master/MALSAR/functions/iMSF
Logi_iMSF	Yuan et al.[49]	https://github.com/jiayuzhou/MALSAR/tree/master/MALSAR/functions/iMSF
iMVWL	Tan et al.[50]	http://mlda.swu.edu.cn/codes.php?name=iMVWL

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Xin Liu (xin3929@163.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Requests for imaging data used in this work should be directed to lead contact. The availability of imaging data will be available upon the specific request, institutional policies, and the project requirements of No.TTHEC-KYSQ 202212401. The accession numbers are also listed in the key resources table.
- The codes are available on a public repository (https://github.com/jq-gogogo/Latent_Relation_Shared_Learning).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The study was conducted in accordance with the Declaration of Helsinki. This study was approved by the Institutional Review Board of Beijing Tiantan Hospital (KYSQ 202212401) and the patient consent was waived.

The experiment involved 117 female(mean age: 55.3 years) native Chinese speakers participants from anonymous patients diagnosed with endometrial cancer through pathology. A total of 93 patients underwent hysterectomy with bilateral salpingo-oophorectomy pelvic and para-aortic lymphadenectomy. An additional 24 patients underwent hysterectomy with bilateral salpingo-oophorectomy and sentinel lymph node biopsy. Postoperatively, 79 patients were diagnosed with stage I endometrial cancer, of which 43 were classified as stage IA and 36 as stage IB.

Additionally, 38 patients were diagnosed with stage II endometrial cancer. All patient information in the database was de-identified, ensuring no subjective impact on the experimental results and excluding any sensitive patient information.

METHOD DETAILS

Main stages of this article can be summarized as follows. (i) we preprocess the patients images by horizontal flipping, vertical flipping, rotating 90, cropping and contrast-modified augmentation methods to increase the diversity of medical images data (see section dataset augmentation); (ii) We propose a deep multi-view fusion network that use MRI images, ultrasound (US) images and contrast-enhanced ultrasound (CEUS) images to identify benign and early-stage endometrial cancer that are difficult to distinguish (see section mapping for multi-modality representation learning); (iii) Optimizing the model so that incomplete multi-modality medical images data can be used for model construction(see section incomplete multi-modality representation learning); (iv) The effectiveness and efficiency of the incomplete multi-modality model are optimized(see section optimization).

Dataset Augmentation

In this work, we collect 117 endometrial cancer samples from anonymous patients diagnosed by pathological, including deep myometrial invasion (DMI) and shallow myometrial invasion (SMI). The size of dataset is shown in Table 5.The groundtruth of this study is lesion confirmation on hospital pathological examination results. We also collect 1228 samples of normal uterine. Each sample consists of three modalities: MRI, ultrasound (US) and contrast-enhanced ultrasound (CEUS) (not all the samples have complete multi-modality data). Our image dataset (TTEC) comes from Beijing Tiantan Hospital, Capital Medical University, and adhered to the principles of the Declaration of Helsinki.

Due to the limited number of samples in the experimental dataset, the model will be overfitted. To ensure the performance of deep learning, data augmentation is applied to increase the medical image diversity. Therefore, horizontal flipping, vertical flipping, cropping, contrast-modified and random noise methods are used to increase the diversity of medical image data, illustrated in Figure 4. After the above processing, the number of images reached 3158.

Mapping for multi-modality representation learning

For image feature extraction, in the framework of latent relation shared representation learning, the latent discriminating feature representation is first learned from multi-modal data in Equation 1, and the transformation from original-to-shared is carried out. In addition, we use the graph Laplacian matrix in Equation 2 and Equation 3 to solve the problem of preserving both global structure features and multi-modality local geometric structure features. We further project multi-label classifier into label-dependent feature space to obtain a comprehensive latent subspace and fully mine consensus feature from each multi-modality data, as detailed in Equation 4.

Let $X^1, X^2, \ldots, X^m$ denotes the original feature matrices derived from multi-modality dataset and $X^v = [x_1^v; x_2^v; \ldots; x_n^v] \in \mathbb{R}^{d\times n}$ as v^{th} view matrix with d dimensions as well as n samples, m indicate the dataset with m views. $Y \in \{0,1\}^{q\times n}$ denotes the corresponding label matrix, where q the number of labels. Hence, $y_i = [y_{1i}; y_{1i}; \ldots; y_{ki}]$ indicate the label vector of the i-th sample. $y_{ji} = 0$ and $y_{ji} = 1$ indicates whether the i-th sample belongs to j-th class.

Since the multi-modality data contains redundant feature information, we assume that there exists latent shared matrices. Therefore, the discriminative low-rank latent subspace model with multi-modality method is given as

$$\min_{V,\{U^v\}} \sum_{v=1}^{m} \|X^v - U^v V\|_F^2 \quad s.t.\ U \geq 0, s.t.\ U \geq 0 \qquad \text{(Equation 1)}$$

where $V \in \mathbb{R}^{k\times n}$ is represented latent feature in these spaces, $U^v \in \mathbb{R}^{d\times n}$ is the so-called basis matrix corresponding to the v^{th} view. k denotes the dimension of the latent subspace. $\|\cdot\|_F$ denotes the Frobenius norm.

Although Equation 1 can excavate the consensus information of multi-modality data, it only has the global structure. To overcome this shortcoming, we use prior learned label-dependent feature space to retain the local geometric structure information. At the same time, theoretically the local geometric structure should also exhibit similar characteristics in the latent subspace V. Therefore, we adopted graph Laplacian matrix to improve information validity of multi-modality local geometric structure:

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\|v_i - v_j\|_2^2 s_{ij} = tr(V(A - S)V^\top) = tr(VLV^\top) \qquad \text{(Equation 2)}$$

where v_i and v_j are the feature representation of the i-th and j-th sample in the latent subspace V, respectively. s_{ij} denotes the similarity between sample i label and sample j label, Equation 3 is detailed. $\top$ denotes the transpose operator. A is a diagonal matrix and $\alpha_{ii} = \sum_{j=1}^{n} s_{ij}$, and S is the affinity graph. So, $L = A - S$ denotes the graph laplacian matrix, which make the retained local feature information more accurate and not to be too drastic with each iteration.

$$s_{ij} = \begin{cases} exp\left(-\frac{\|y_i - y_j\|_2^2}{\sigma^2}\right), y_i \in \mathcal{N}_p(y_j) \text{ or } y_j \in \mathcal{N}_p(y_i) \\ 0, otherwise \end{cases} \qquad \text{(Equation 3)}$$

where $\mathcal{N}_p(y_j)$ denotes the p-nearest neighbors of instance y_i, σ denotes the regularization parameter. Nearest neighbor graph on the label Y is used to construct the local geometric structure.

The function of Equation 2 embedded into Equation 1 is to incorporate prior label information into subspace learning, thereby obtaining the required label-dependent feature space. In order to achieve the goal of disease classification, our multi-label classifier is introduced to map the learned label-dependent feature space to label space. This could be present as:

$$\min_{V,\{U^v\},W} \sum_{v=1}^{m} \|X^v - U^v V\|_F^2 + \alpha \|WV - Y\|_F^2 + \beta tr(VLV^\top) + \gamma \|W\|_F^2 \quad \text{(Equation 4)}$$

where α, β, γ are trade-off parameters that is used to balance the reconstruction error. $W \in \mathbb{R}^{q \times k}$ is represented the prediction learning model. The goal is to give enough weight to informative features and as little weight as possible to less informative features. In Equation 4, with the help of the graph laplacian matrix, we can obtain a comprehensive latent subspace to fully mine the consensus information from each multi-modality data.

So far, the method in Equation 4 can only be applied to fully modal data, i.e., each patient has the MRI, US and CEUS. Patients with incomplete data will be discarded.

Incomplete multi-modality representation learning

Equation 4 can only be applied to the complete multi-modality image data, so incomplete multi-modality data will be discarded. In fact, there are many useful information in incomplete data. Furthermore, medical image data are often very limited, and if these data are discarded directly, the performance of the model will be affected. We refined on Equation 4 to make use of these incomplete multi-modality patient image data.

Let $X^v = [X^{vc}; X^{v\bar{c}}] \in \mathbb{R}^{d \times (n^c + n_o^{\bar{c}})}$ denotes the original feature matrices of v^{th} modality data, where $n^c + n_o^{\bar{c}}$ is the total number of patients(n^c represents the number of patients with complete multi-modality data, and $n_o^{\bar{c}}$ represents the number of patients with o-th modality data). Certain special cases need to be taken into consideration. For example, if samples with only two modalities are taken, it would be transformed into complete multi-modality problems, in which case Equation 4 can be applied accordingly. Similar, $Y^v = [Y^{vc}; Y^{v\bar{c}}]$ and $U^v = [U^{vc}; U^{v\bar{c}}]$. Using the notations above, we summarized as follows:

$$\mathrm{Min}_{V,\{U^v\},W} \sum_{v=1}^{m} \|[X^{vc}; X^{v\bar{c}}] - [U^{vc}; U^{v\bar{c}}]V\|_F^2 + \alpha \|WV - [Y^{vc}; Y^{v\bar{c}}]\|_F^2 + \beta tr(VLV^\top) + \gamma \|W\|_F^2 \quad \text{(Equation 5)}$$

In this way, we can use as much limited patient data as possible to participate the model training. In addition, no other noise parameters are introduced in this method, and we don't have to discard incomplete patient data, which discarding limited patient data are more likely lead to overfitting.

Optimization

To address the optimization problem of Equation 5, which is not jointly convex with respect to V, W, U^v coupled together. We take an alternative optimizer approach to optimize the variables to obtain a local minimum.

Updating V: Fix W and U^v, then update V:

$$\min_{V} \sum_{v=1}^{m} \|[X^{vc}; X^{v\bar{c}}] - [U^{vc}; U^{v\bar{c}}]V\|_F^2 + \alpha \|WV - [Y^{vc}; Y^{v\bar{c}}]\|_F^2 + \beta tr(VLV^\top) \quad \text{(Equation 6)}$$

Taking derivative of V and making it to 0,

$$\left(\sum_{v=1}^{m} 2[U^{vc}; U^{v\bar{c}}]^\top [U^{vc}; U^{v\bar{c}}] + 2\alpha V^\top V \right) V + V\beta(L + L^\top) = \sum_{v=1}^{m} [U^{vc}; U^{v\bar{c}}]^\top [X^{vc}; X^{v\bar{c}}] + 2\alpha V^\top [Y^{vc}; Y^{v\bar{c}}] \quad \text{(Equation 7)}$$

Define A, B, and C as fellows,

$$\begin{cases} A = \sum_{v=1}^{m} 2[U^{vc}; U^{v\bar{c}}]^\top [U^{vc}; U^{v\bar{c}}] + 2\alpha V^\top V \\ B = V\beta(L + L^\top) \\ C = \sum_{v=1}^{m} [U^{vc}; U^{v\bar{c}}]^\top [X^{vc}; X^{v\bar{c}}] + 2\alpha V^\top [Y^{vc}; Y^{v\bar{c}}] \end{cases} \quad \text{(Equation 8)}$$

Then the Equation 7 can be to a Sylvester equation problem, which can be easily solved.

$$AV + VB = C \quad \text{(Equation 9)}$$

Algorithm 1. Optimization

Require: m views $\{X^1, X^2, \ldots, X^m, Y\}$, $X^v \in \mathbb{R}^{d \times n}$, $Y \in \mathbb{R}^{q \times n}$

Ensure: Multi-label classifier W.

1: randomly initialize V, W and U^v, the nearest neighbor number $p = 20$

Repeat:

2: updating V by Equation 9

3: updating W by Equation 10

4: updating U^v by Equation 12

Until:

5: reach the maximum iteration or converged.

End for:

6: Multi-label classifier **W**.

Updating W: Fix V and U^v, the optimization problem becomes,

$$\min_{W} \alpha \|WV - [Y^{vc}; Y^{v\bar{c}}]\|_F^2 + \gamma \|W\|_F^2 \quad \text{(Equation 10)}$$

Then, W is updated as $W = \frac{YV^\top}{VV^\top + \frac{\gamma}{\alpha} I}$, $I \in \mathbb{R}^{k \times k}$ is an identity matrix.

Updating U^v: Fix V and W, the optimization problem becomes,

$$\min_{\{U^v\}} \|[X^{vc}; X^{v\bar{c}}] - [U^{vc}; U^{v\bar{c}}]V\|_F^2 \quad \text{(Equation 11)}$$

Taking derivative of U^v and making it to 0,

$$U^v = [X^{vc}; X^{v\bar{c}}]V^\top (VV^\top + \eta I) \quad \text{(Equation 12)}$$

where η denotes a small regularization parameter.

After updating the V, W and U^v, we can get the optimal solution. Algorithm 1 presents the optimization part.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were performed using custom python scripts. Details of all statistical analyses can be found above in the relevant subsections of the method details section. Continuous variables are given in the form of means ± standard deviations while categorical variables are presented as numbers (%).