Chinese Medical Journal®

Development and validation of deep learning for predicting the growth of ovarian cancer organoids

Hongji Wu[1], Lifang Ma[2,3,4], Ling Wang[2,3,4], Xueping Zhu[2,3,4], Xiaogang Luo[1], Cong Zhang[5], Chunfang Ha[6], Yun Dang[7,8], Haixia Wang[2,3,4], Dongling Zou[2,3,4]

[1]Bioengineering College, Chongqing University, Chongqing 400044, China;
[2]Department of Gynecologic Oncology, Chongqing University Cancer Hospital & Chongqing Cancer Institute & Chongqing Cancer Hospital, Chongqing 400030, China;
[3]Chongqing Specialized Medical Research Center of Ovarian Cancer, Chongqing 400030, China;
[4]Organoid Transformational Research Center, Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing 400030, China;
[5]School of Medicine, Chongqing University, Chongqing 400044, China;
[6]Gynecology Department, General Hospital of Ningxia Medical University, Yinchuan, Ningxia 750003, China;
[7]Gansu Provincial Maternity and Child Care Hospital, Lanzhou, Gansu 730050, China;
[8]Institute of Pathology, School of Basic Medical Sciences, Lanzhou University, Lanzhou, Gansu 730000, China.

Abstract

Background: Organoids have attracted enormous interest in disease modeling, drug screening, and precision medicine. However, developing robust patient-derived organoids (PDOs) was time-consuming, costly, and had low success rates for certain cancer types, which limited their clinical utility. This study aimed to develop an interpretable deep learning-based model to predict the cultivation outcome of ovarian cancer organoids in advance.

Methods: Longitudinal microscopy images of 517 ovarian cancer organoid droplets were divided into training ($n = 325$), validation ($n = 88$), and test ($n = 104$) sets. Subsequently, growth prediction models were developed based on four neural network backbones (ResNet18, VGG11, ConvNeXt v2, and Swin Transformer v2), and specific optimization methods were designed for better prediction. Finally, 179 samples from multiple centers were collected for prospective validation, and the gradient-weighted class activation mapping (Grad-CAM) method was used for interpretability analysis of the deep model to reveal the basis of the model's decisions.

Results: The test set showed that the deep learning models could achieve high-performance prediction at the third stage with area under the curve (AUC) values greater than 0.8 for all four models. The homogeneous transfer learning optimization method improved the AUC from 0.833 to 0.884 ($P = 0.0039$). In prospective validation, the optimized model achieved an AUC of 0.832, a Brier score of 0.1919 in the calibration curve, and a greater net benefit in the decision curve. Interpretability analysis revealed that the area where organoids are being formed and have already formed is important for prediction.

Conclusions: Our developed models achieved satisfactory results in predicting the growth of ovarian cancer organoids. There is potential for further development of the model toward process automation.

Keywords: Organoid; Artificial intelligence; Deep learning; Ovarian cancer; Precision medicine

Introduction

Organoids are three-dimensional structures composed of various types of self-organizing stem cells that can simulate the structure and function of native organs. Therefore, they are helpful for modeling human diseases, drug screening, and *in vitro* regeneration of alternative tissues or organs. Compared to patient-derived xenografts, organoids have advantages such as short culture cycles and easy operation.[1] Unlike traditional two-dimensional (2D) cell line cultures, organoids are directly derived from patients, can preserve the genomic and epigenetic information of their donors, and have a three-dimensional (3D) structure that

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.1097/CM9.0000000000003575

Hongji Wu and Lifang Ma contributed equally to this work.

Correspondence to: Haixia Wang, Department of Gynecologic Oncology, Chongqing Specialized Medical Research Center of Ovarian Cancer, Organoid Transformational Research Center, Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing 400030, China
E-Mail: wanghx1985@126.com;
Dongling Zou, Department of Gynecologic Oncology, Chongqing Specialized Medical Research Center of Ovarian Cancer, Organoid Transformational Research Center, Chongqing Key Laboratory of Translational Research for Cancer Metastasis and Individualized Treatment, Chongqing University Cancer Hospital, Chongqing 400030, China
E-Mail: cqzl_zdl@163.com

Copyright © 2025 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2026;139(1)

Received: 20-11-2024; **Online:** 25-07-2025 **Edited by:** Yanjie Yin

is more in line with those encountered *in vivo* conditions. They can more accurately summarize the structure, specific functions, molecular characteristics, and genomic changes. Organoids have great potential in predicting treatment responses, studying drug resistance-related mechanisms, optimizing treatment strategies, and enabling precision medicine. After years of development, scientists have successfully established various cancer organoid models, such as prostate cancer,[2] colorectal cancer,[3] pancreatic cancer,[4] and ovarian cancer models.[5]

However, there are still many limitations in organoid-guided precision medicine. First, 2–4 weeks are usually necessary for culturing patient-derived organoids (PDOs). Second, the establishment rates for some types of cancers remain low. It has been reported that the success rate for constructing high-grade serous ovarian cancer organoids is only 23–53%.[6] Third, the cultivation process incurs high costs due to the necessity of using expensive growth factors and other materials.[7] These findings indicate that developing ovarian cancer organoids is highly likely to fail, leading to a waste of time and money. If the cultivation outcomes of organoids can be predicted early in their growth, allowing for the early termination of unsuccessful samples and avoiding ineffective cultivation, this can promote the efficient advancement of ovarian cancer organoid research and clinical translation.

Artificial intelligence, particularly deep learning, has significant advantages in automated feature extraction and large-scale data processing, which has allowed it to shine in medical fields such as lesion detection, disease diagnosis, and drug discovery.[8–10] The integration of deep learning techniques with diverse data types, such as imaging, clinical, and genomic data, has greatly enhanced the accuracy, efficiency, and personalization of medical diagnoses, while also promoting the scientific and precise nature of clinical decision-making.[11–14] Given the notable efficacy of deep learning in biology and medicine, more organoid researchers are using technologies and methods by integrating deep learning to improve and automate work. Yang *et al*[15] recovered the phase of organoid images using a deep learning-based method that does not require a training set to obtain high-quality organoid images. Kassis *et al*[16] proposed a dataset of manually annotated human intestinal organoid images and a deep convolutional neural network called OrgaQuant for automatic detection and counting of organoids in brightfield images. Bian *et al*[17] trained Single Shot Detector (SSD) model to detect all organoids and designed a model based on ResNet to dynamically track each organoid throughout the entire culture process. Abdul *et al*[18] developed a deep learning-based imaging tool to automatically assess the crypt formation and opacity of colorectal organoids from brightfield images and determine the extent of organoid structural maturity for the classification of organoid morphology. Lim *et al*[19] predicted biomarker expression based on brightfield images of organoids using a convolution neural network for assessing airway organoid differentiation. Larsen *et al*[20] proposed a model called the regularized conditional adversarial network to generate corresponding dye fluorescence images and overall viability from brightfield images for drug assays.

Inspired by the success of deep learning in organoid image analysis, we hypothesized that brightfield images contain sufficient information and can be extracted using deep learning to predict organoid growth. Most similar to our research, Kegeles *et al*[10] developed a deep learning-based computer algorithm to recognize and predict retinal differentiation in stem cell-derived organoids. Notably, each of their organoids was seeded in a separate well, which is different from the procedure used in ovarian cancer organoid culture, which adopts bulk culture methods. Organoids seeded close to other organoids may be affected by cell–cell paracrine or juxtacrine signaling.[21] In addition, analysis of organoid images at a single, subjectively selected time may miss the optimal prediction window.

The purpose of this study was to develop and validate deep learning-based prediction models for ovarian cancer organoid growth to identify which ovarian cancer organoid droplets will grow successfully as early as possible. First, since there is no public dataset available for this task, this study constructed a high-throughput ovarian cancer organoid image dataset that records the dynamic process of organoid culture. Second, prediction models for ovarian cancer organoid growth were developed based on four deep learning models, achieving high-performance prediction of cultivation outcomes at an early stage of ovarian cancer organoid growth. In addition, this study proposed a transfer learning method to make the model obtain prior knowledge for better prediction and the effectiveness of this method was confirmed on the test set and through multicenter prospective validation. Third, interpretability analysis based on gradient-weighted class activation mapping (Grad-CAM) enhanced confidence in the further application of deep learning in organoids.

Methods

Study design and samples

As shown in Figure 1, this study consisted of three phases to train and comprehensively validate deep learning models for predicting ovarian cancer organoid growth. In phase 1, ovarian cancer organoid samples were collected from Center 1 and randomly divided into training, validation, and test sets. Numerous data augmentations were implemented to augment the training set. In phase 2, four selected deep learning models were trained using the training and validation sets to construct deep learning models for prediction at different time points and using different image magnifications. A comparison was made on the test set to determine the optimal prediction time and image magnification. Subsequently, specialized optimization methods for this task were proposed and assessed by the area under the curve (AUC), calibration curve, and decision curve to determine improvement. In phase 3, prospective validation was conducted using the optimized models. Finally, interpretability analysis was performed on the model to reveal its decision basis.

A total of 696 ovarian cancer organoid droplet samples from three research institutes were included in this study. In phase 1, 517 organoid droplet samples were collected

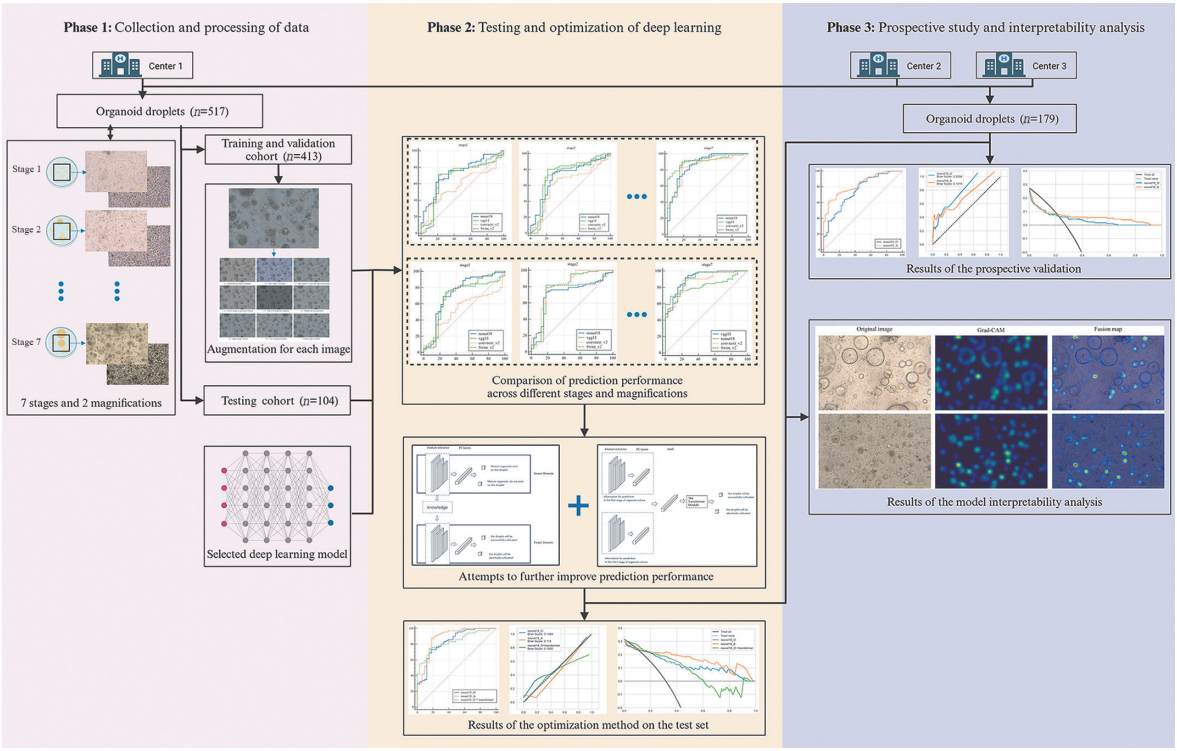


Figure 1: Framework from culture to prediction. The study includes three phases: data collection and processing of organoid droplets across 7 stages (Phase 1), training and optimization of deep learning models with stage and magnification comparisons (Phase 2), and prospective validation with interpretability analysis using Grad-CAM (Phase 3). Grad-CAM: Gradient-weighted class activation mapping.

from April 2022 to June 2023 at the Organoid Transformational Research Center, Chongqing University Cancer Hospital. Prospective experiments started in July 2023 and ended in February 2024. A total of 179 samples were included from the Organoid Transformational Research Center at Chongqing University Cancer Hospital, the Department of Gynecology and Oncology at Gansu Provincial Maternity and Child Care Hospital, and the Department of Gynecology at the General Hospital of Ningxia Medical University. In accordance with the *Declaration of Helsinki*, all patient specimens were obtained with the approval of the the Ethics Committee of Chongqing University Cancer Hospital (No. CZLS2023151-A), Ningxia Medical University General Hospital Scientific Research Ethics Committee (No. KYLL-2022-1236) and written informed consent.

Image capture and preprocessing

Image capture began one or two days after seeding. The process involves several key steps: (1) the objective lens magnification was set to 4 times; (2) the *Z*-stacking function was used to synthesize images of the droplet at different focal planes where cells exist; and (3) the objective lens magnification was adjusted to 10 times, and the same capturing procedure was repeated. After the first image was captured, each two-day interval constitutes one stage, resulting in seven consecutive stages being captured. The image resolution was 2736 px × 1824 px.

Before feeding the training images to the models, to simulate the different operators and shooting conditions and increase the training set size, numerous data augmentations were implemented: (1) horizontal and vertical flipping, (2) motion blurring, (3) brightness and saturation changes, (4) image conversion to grayscale versions, (5) addition to the brightness channels, (6) hue and saturation changes, (7) image contrast adjustments, (8) horizontal flipping, and (9) vertical flipping.

Base models

For our task, we selected the following models with different architectures that were effective at ImageNet competitions and in multiple biological applications: VGG11,[22] ResNet18,[23] ConvNeXt v2,[24] and Swin Transformer v2.[25]

To enhance the generalization ability of the model, we used the ImageNet database to pretrain four models and then fine-tuned our training dataset. This can increase the convergence speed of the model and suppress the influence of overfitting.

In this study, we chose the stochastic gradient descent (SGD) optimizer.[26] We developed the model training procedure based on our own dataset and the learning rate adjustment strategy proposed in the original paper for four models. For the hyperparameters, we adjusted the learning rate and corresponding iteration epochs to ensure that the models performed well on both the training and validation datasets. The final hyperparameters are shown in Supplementary Table 1, http://links.lww.com/CM9/C395. The training of the deep

learning models for prediction at different time points and using different magnification images is illustrated in Supplementary Figure 1, http://links.lww.com/CM9/C395. No significant underfitting or overfitting is observed.

We implemented the proposed deep learning models with the PyTorch network framework and trained the models on a GTX3090TI GPU.

Model optimization

To improve the prediction performance, we used two optimization methods: (1) We selected photos with mature organoids and almost nonexistent organoids in the training set to train a binary classification model capable of identifying the presence of ovarian cancer organoids in an image. Subsequently, fine-tuning from scratch[27] was performed [Supplementary Figure 2A, http://links.lww.com/CM9/C395]. (2) The architecture of combining ResNet18 whose parameters are frozen with Transformer[28] was adopted to capture the temporal information rather than just the spatial information of two different stages of images of organoid droplets [Supplementary Figure 2B, http://links.lww.com/CM9/C395].

Interpretation of the model

Deep learning has improved biomedical image analysis. However, it is viewed as a black box that does not explain how predictions are made, which has led to relatively low acceptance of deep learning among healthcare practitioners. Therefore, an explanation and interpretation of the model's output and work are required to improve the use of these systems in various clinical applications.[29]

This study used Grad-CAM[30] to produce visual explanations for growth prediction decisions from deep learning models. Specifically, after feeding an image into the well-trained convolutional neural network (CNN) model, it extracts feature maps with multiple channels from the final convolutional layer through forward propagation. Then, the gradient weights, which contain the importance of each channel, were obtained by using the final prediction score to calculate the gradient information and move it back to the final convolutional layer through back-propagation. Then, the feature maps and gradient weights were multiplied to generate the weighted combination of feature maps.[31]

Statistical analysis

Codes written in Python (https://www.python.org), the imgaug library (version 0.4.0), Torchvision (version 0.14.0), PyTorch (version 1.13.0), Pandas (version 1.5.2), skimage (version 0.19.3), and Matplotlib (version 3.6.3) were used for image processing, model development and training, data analysis, and visualization. The AUCs were analyzed and compared using MedCalc software 22 (MedCalc Software Ltd, Ostend, Belgium). The 95% confidence intervals (CIs) for the models' AUCs were calculated using the binomial exact method. Comparisons between different models were conducted using the DeLong test (P <0.05 indicated statistical significance).

Results

Ovarian cancer organoid dataset and preprocessing

In this paper, we created the first dynamic high-throughput ovarian cancer organoid dataset for automatic growth prediction of organoids. The dataset contains sampling data for 696 organoid droplets. It is composed of different types of ovarian cancer [Figure 2A]: high-grade serous, low-grade serous, mucinous, endometrioid ovarian cancer, and clear cell cancer, with the focus mainly on high-grade serous cancer. The outcome of organoid droplet was annotated by professionals, and typical successful and failed organoid droplets are shown in Figure 2B.

Due to the rarity of brightfield images of ovarian cancer organoids, we enriched the training dataset before deep learning training. We applied channel shuffling, motion blurring, flipping, and other methods for data augmentation [Figure 2C] to imitate different microscope operators and shooting conditions. As a result, 10-fold more data than the initial bright-field images were used to train the models.

Comparison of performance at different magnifications and stages

To analyze the optimal magnification and prediction time, this study used microscopy images at different magnifications and stages to train the models and then compared the results on the test set. The prediction results using 4× or 10× images at different time points are shown in Table 1.

When comparing different magnifications, as shown in Table 2, the P value of the AUC comparison indicates that predictions using 10× images were either significantly greater or showed no significant difference compared to those using 4× images at various stages, with no instances of significantly lower performance. Compared with 4× magnification, 10× objective lens magnification is more suitable for predicting the growth of ovarian cancer organoids.

For the optimal prediction time, as shown in Table 1 and Supplementary Figure 3, http://links.lww.com/CM9/C395, when using 10× images in the third stage for prediction, the AUC achieved excellent results: all four models reached AUCs of over 0.8 (0.869, 0.833, 0.855, and 0.876). In addition, as shown in Table 3, the P value of the AUC comparison between different stages and the previous stage indicates that predictions using stage-3 images significantly improved compared to those using stage-2 images, whereas there was no significant improvement from stage-3 to stage-4 images. Considering the prediction efficiency and performance, this study concluded that the third stage of ovarian cancer organoid droplet culture is the optimal prediction time.

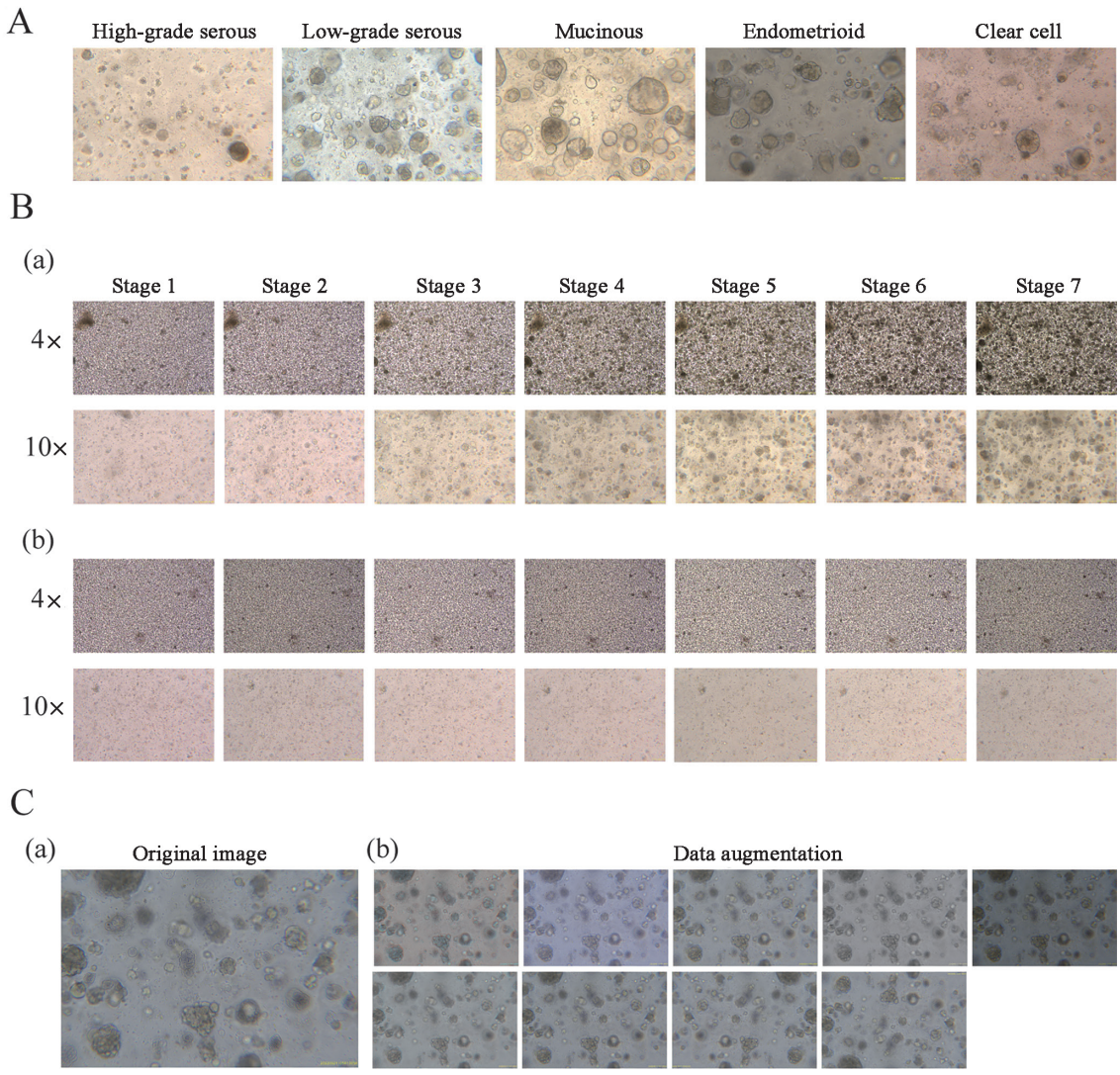


Figure 2: Ovarian cancer organoid dataset and preprocessing. (A) Different pathological types of ovarian organoids (Original magnification ×10). (B) Images of typical successful and failed organoid droplets at seven different time points. (a) Brightfield images at 4× and 10× magnifications of a successfully cultivated representative organoid droplet, as captured at seven different time points. (b) Brightfield images at 4× and 10× magnifications of a representative organoid droplet with failed cultivation, which were captured at seven different time points. (C) Original and augmented images of organoid droplets (Original magnification ×10). (a) Original image. (b) Data augmentation techniques simulating different photographic environments and photographers.

Table 1: Area under the curves (95% confidence intervals) of the four models using 4× or 10× images in the test cohort.

		Stage						
Model	**Magnification**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
ResNet18	4×	0.732 (0.636–0.814)	0.719 (0.622–0.803)	0.831 (0.745–0.898)	0.759 (0.665–0.837)	0.814 (0.726–0.884)	0.784 (0.692–0.858)	0.867 (0.787–0.926)
	10×	0.730 (0.634–0.812)	0.723 (0.627–0.806)	0.833 (0.747–0.899)	0.842 (0.758–0.906)	0.769 (0.676–0.846)	0.874 (0.795–0.931)	0.825 (0.738–0.893)
VGG11	4×	0.679 (0.581–0.767)	0.747 (0.653–0.827)	0.828 (0.742–0.895)	0.765 (0.671–0.842)	0.822 (0.735–0.890)	0.783 (0.691–0.857)	0.911 (0.839–0.958)
	10×	0.678 (0.579–0.766)	0.787 (0.695–0.861)	0.869 (0.788–0.927)	0.904 (0.830–0.953)	0.830 (0.743–0.896)	0.914 (0.842–0.960)	0.909 (0.837–0.957)
ConvNeXt v2	4×	0.588 (0.487–0.684)	0.630 (0.529–0.722)	0.696 (0.598–0.782)	0.772 (0.679–0.849)	0.783 (0.692–0.858)	0.779 (0.687–0.855)	0.848 (0.764–0.911)
	10×	0.530 (0.429–0.628)	0.817 (0.729–0.886)	0.855 (0.772–0.916)	0.869 (0.789–0.927)	0.877 (0.798–0.933)	0.945 (0.882–0.980)	0.877 (0.798–0.933)
Swin v2	4×	0.681 (0.583–0.769)	0.720 (0.624–0.804)	0.848 (0.765–0.911)	0.778 (0.686–0.854)	0.790 (0.700–0.864)	0.760 (0.666–0.838)	0.883 (0.805–0.938)
	10×	0.712 (0.615–0.797)	0.730 (0.634–0.812)	0.876 (0.797–0.933)	0.850 (0.767–0.913)	0.830 (0.744–0.897)	0.832 (0.747–0.899)	0.892 (0.816–0.944)

Table 2: Comparison of the area under the curves for the deep learning models on the testing set using 10× images and 4× images.

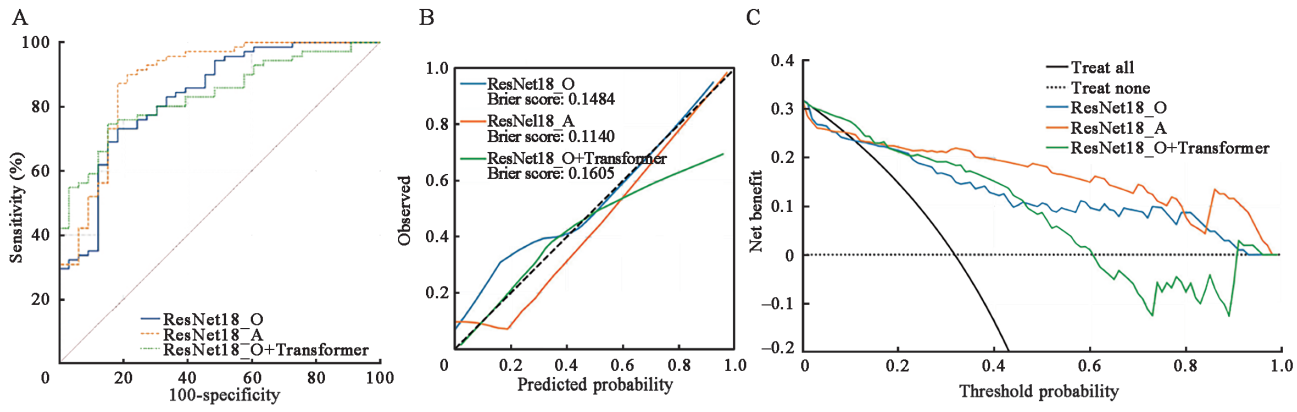
	ResNet18		VGG11		ConvNeXt v2		Swin v2	
Stages	*Z* value	*P* value	*Z* value	*P* value	*Z* value	*P* value	*Z* value	*P* value
1	0.0680	0.9458	0.0610	0.9514	1.7546	0.0793	1.4937	0.1352
2	0.1359	0.8920	1.2875	0.1980	5.1498	<0.0001	0.6362	0.5247
3	0.0514	0.9590	1.7191	0.0856	5.2455	<0.0001	1.5585	0.1191
4	2.6024	0.0093	4.0852	<0.0001	3.3431	0.0008	2.8380	0.0045
5	1.5930	0.1112	0.4501	0.6526	2.8234	0.0048	1.7431	0.0813
6	2.4862	0.0129	3.9436	<0.0001	4.1719	<0.0001	2.7509	0.0059
7	1.0668	0.2860	0.0899	0.9280	1.3318	0.1829	0.4643	0.6424

Table 3: Comparison of the area under the curves of the deep learning models on the testing set, with predictions made with 10× images at different time points.

	ResNet18		VGG11		ConvNeXt v2		Swin v2	
Stage	*Z* value	*P* value	*Z* value	*P* value	*Z* value	*P* value	*Z* value	*P* value
2 *vs.* 1	0.1834	0.8545	4.2296	<0.0001	6.4671	<0.0001	0.7135	0.4756
3 *vs.* 2	3.9599	<0.0001	2.8602	0.0042	1.9379	0.0526	4.6775	<0.0001
4 *vs.* 3	0.3603	0.7186	1.4413	0.1495	0.7857	0.4320	1.1693	0.2423

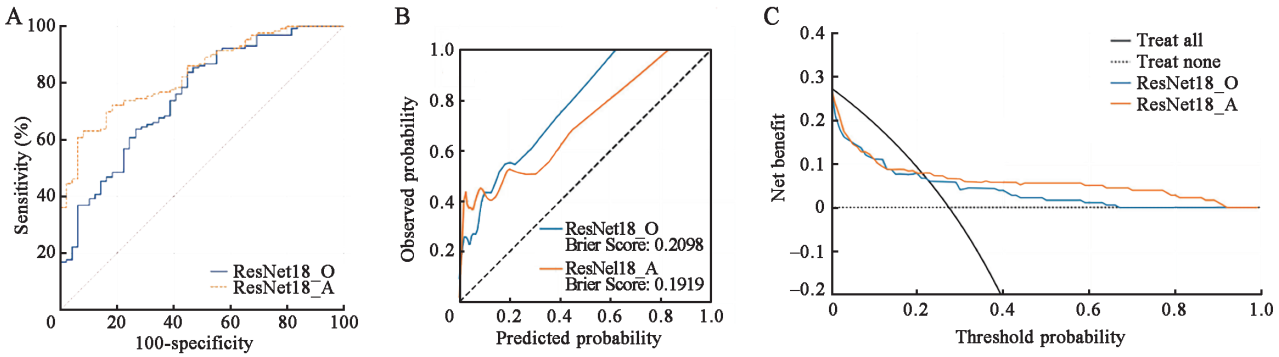
Comparison of performance before and after optimization

For the optimization attempts, the performance using 10× images in the third stage of the test set is shown in Figure 3. The AUC comparison and receiver operator characteristic (ROC) curves are presented in Figure 3A. Compared with the original ResNet18 (ResNet18_O), the model fine-tuned after classification (ResNet18_A) has a positive effect on the prediction performance ($z = 2.887$, P value = 0.0039). The second attempt (ResNet18_O+Transformer) had a greater AUC than did ResNet18_O, but the improvement was not statistically significant ($z = 0.109$, P value = 0.9128). The calibration curves for predictions at the third time point are depicted in Figure 3B. ResNet18_A exhibited the smallest Brier score (0.1140 *vs.* 0.1484 *vs.* 0.1605). For the decision curves of the models in the test set, as shown in Figure 3C, ResNet18_A demonstrated a clear advantage. Both ResNet18_A and ResNet18_O yielded positive net benefits except for low and extremely high thresholds, with the former generally outperforming the latter and resulting in greater positive net benefits. ResNet18_O+Transformer only generates positive net benefits between low and relatively high thresholds. Overall, ResNet18_A exhibits better performance on the test set. The proposed homogeneous transfer learning optimization approach effectively enhances the predictive performance of deep learning models for ovarian cancer growth.



Model	AUC (95% CI)	*Z* value	*P*-value
ResNet18_O	0.833 (0.747–0.899)		
ResNet18_A	0.884 (0.807–0.939)	2.887	0.0039
ResNet18_O+Transformer	0.835 (0.750–0.901)	0.109	0.9128

Figure 3: Performance of the model before and after optimization on the test set. (A) Receiver operator characteristic curves and corresponding area under the curves. (B) Calibration curves. (C) Decision curves. CI: Confidence interval.



model	AUC (95% CI)	*Z* value	*P*-value
ResNet18_O	0.755 (0.685–0.816)		
ResNet18_A	0.832 (0.769–0.884)	3.001	0.0027

Figure 4: Performance of the model in the prospective validation and interpretability analysis. (A) Receiver operator characteristic curves and corresponding area under the curve of the models in the prospective validation. (B) Calibration curves. (C) Decision curves. CI: Confidence interval.

Prospective validation and interpretation of the model

ResNet18_A, which was chosen for its outstanding performance on the test set, was used for prospective validation in this study. Moreover, ResNet18_O was also used for prospective validation to contrast with A to investigate whether the optimization effectively enhanced generalization. Predictions were made in the third stage before the cultivation outcome was known. The performance of the models in the prospective validation showed similar results. The AUC comparison and ROC curves are presented in Figure 4A. ResNet18_A still outperformed ResNet18_O, with an AUC of 0.832 compared to 0.755. The calibration curves for the prospective experiment results are depicted in Figure 4B, with Brier scores of 0.2098 for ResNet18_O and 0.1919 for ResNet18_A. However, the curves are above the reference line, indicating that the models underestimated the risk of cultivation failure in the prospective validation. The decision curves for the prospective experiment results are shown in Figure 4C, where ResNet18_A demonstrates positive net benefits across most thresholds. On the other hand, the threshold range for ResNet18_O to achieve positive net benefits is relatively narrow. Moreover, ResNet18_A continues to exhibit excellent performance as a classifier in prospective validation.

Additionally, this study used Grad-CAM to analyze the decision-making basis of the model, further verifying its reliability. Grad-CAM computes the gradients of the weights associated with distinguishable features, and its final output is a heatmap, highlighting the regions the model sees as important for its final prediction.[19] In the heatmap, the red regions contribute the most, whereas the blue regions are less important. For ResNet18_A in the third stage, the area where organoids are forming and have already formed is important for predicting whether organoid droplets can be successfully generated [Figure 5A]. For samples that failed to cultivate, Grad-CAM showed little to no corresponding regions, supporting its ability to predict successful cultivation [Figure 5B]. This reveals that the image regions that deep learning

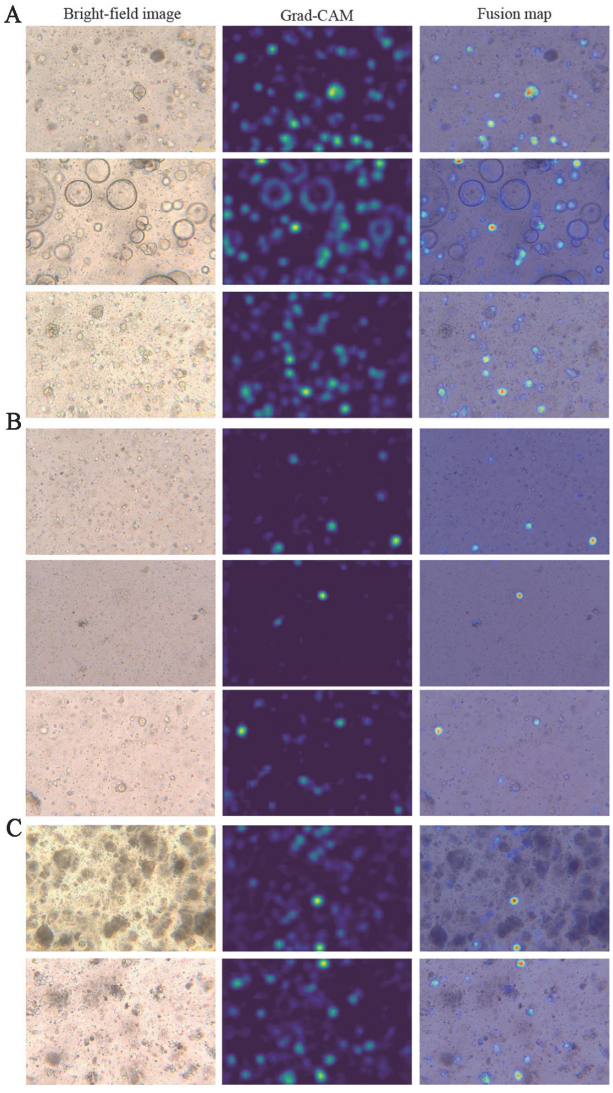


Figure 5: Interpretability analysis of the model predictions via Grad-CAM on some 10× samples. (A) Ovarian cancer organoid droplets correctly predicted to be successfully cultivated by the model in the third stage. (B) Samples correctly predicted to fail in cultivation by the model. (C) Sample was incorrectly predicted by the model to be successfully cultivated in the third stage. Grad-CAM: Gradient-weighted class activation mapping.

models focus on for prediction are consistent with those of experts. However, we also identified a pattern that was not correctly discerned: There were numerous cell clusters, but some had begun to disintegrate or had already disintegrated, which often led to incorrect predictions of successful cultivation. A typical sample is Figure 5C, which the model does not correctly predict.

Discussion

Improving the efficiency of organoid cultivation in organoid research has attracted much attention. Artificial intelligence is now a powerful tool in medical analysis. Artificial intelligence has successfully used patient image information to detect diabetic retinopathy[11] and predict tumor growth.[14] This has promoted the automation and intelligence of health care, accelerating the timeline of diagnosis and treatment. Current applications of artificial intelligence in organoid research[16,18,20,32,33] primarily focus on immediate metrics of organoids, such as identification, counting, and classification, without the ability to predict the fate of organoids in advance. Although it is difficult to predict cultivation outcomes based on early morphological features of organoids to terminate potentially failed samples in advance, such predictions are traceable. Previous studies revealed that cell fate determination is critically influenced by changes in 3D tissue architecture, cellular rearrangements, cytoskeletal flow, etc.,[34] showing that spatial organization influences temporal progression.[35] The combination of morphological analysis and artificial intelligence technology has made significant progress in predicting cell development[36–38] and differentiation.[39–41] Therefore, we propose the use of artificial intelligence technology combined with brightfield microscopy images of ovarian cancer organoids to predict the growth outcome of organoids in the early stages of cultivation. Many design decisions involved in achieving this prediction task require the input of human experts, including data processing, model training and optimization, and comprehensive validation.

Artificial intelligence can accomplish classification tasks through two distinct approaches. One is traditional machine learning, characterized by its strong reliance on feature engineering,[42] which requires manually extracting and combining effective features to provide the necessary inputs for the model. The other approach is deep learning, known for its end-to-end learning capability,[9,43] which allows models to automatically learn data representations with minimal human intervention. There is a lack of mature and precise features and standards for evaluating the developmental potential of ovarian cancer organoids, so this study adopted a deep learning approach. Here, we showed that such classification is possible with deep learning, which can predict early ovarian cancer organoid growth based only on brightfield images of ovarian cancer organoid droplets.

Deep learning is a data-driven algorithm model that requires a large amount of training data to optimize parameters.[17] Since there is no public dataset available for our task, we created a dynamic high-throughput ovarian cancer organoid image dataset that spans more than two weeks. We constructed four growth prediction models for ovarian cancer organoids using deep learning. However, considering that data scarcity is a core issue in obtaining medical images and applying machine learning techniques to train deep learning models for medical applications,[44] we additionally used two widely regarded effective strategies to address this challenge.[45–47] One is to use a network with parameters pretrained on a large-scale image dataset such as ImageNet as the backbone and fine-tune the network on our specific dataset. The other is to use data augmentation methods to simulate imaging under various shooting conditions and by different photographers, thereby increasing the number and diversity of training samples.

Four deep learning models with outstanding performance in large-scale classification datasets, including both classical models and cutting-edge models, were selected for this study. These models have been adopted in many automatic and reliable computer-aided diagnosis systems for medicine, such as brain tumor classification[48] and skin disease diagnosis.[49] The four models exhibited excellent performance on the test set after being trained on the training set in this study, and when using images from the third stage of cultivation (fifth or sixth day) for prediction, the AUC of the four models reached approximately 0.85. This shows the possibility of combining deep learning with organoid brightfield images to predict organoid growth in the early stages. According to the experimental results, the growth of most ovarian cancer organoid droplets can be accurately predicted within one week. This study can be integrated with automated lab systems, enabling end-to-end objective monitoring by automatically predicting culture outcomes based on data from automated microscopy; providing developmental trend information to help the culture system dynamically adjust conditions with personalization; and incorporating other data collected by the system for in-depth analysis of relevant biological processes, thereby reducing the time and cost of organoid culture and accelerating its automation and clinical translation.

To further improve the prediction performance, we attempted two specific optimization methods tailored to the task characteristics. The first method proposed a homogeneous transfer learning approach to provide prior knowledge to guide the learning of prediction tasks.[50] Although many medical deep learning applications directly train pretrained models using ImageNet data with their own datasets, this approach may not be effective considering the distance between the source and target domains.[51] The transfer learning approach proposed in this study enables good overlap between the feature spaces of the source and target domains because the two tasks are correlated to some degree (if there are mature organoids in the early stages in a sample, it is clearly more likely to succeed). In our experiment, the AUC increased by 5.1% when the method was used. The second method aimed to make the deep learning model learn not only the spatial features of the droplets but also the temporal information of the droplets. The temporal nature of the data, such as the ability to predict organoid ATP from images, is an important factor that can improve

organoid analysis.[52] However, experiments showed that it did not have a significant effect. This method may introduce more irrelevant noise information and increase the dimensionality of features,[14] although it can introduce temporal information. Therefore, we opted for the first optimization method. In the prospective validation, the AUC, calibration curve, and decision curve also indicated the effective enhancement of the model's generalizability.

However, deep learning models require interpretability. In previous studies, researchers created models that appeared to perform well on their test sets yet, upon further inspection, based their decisions on confounding information (e.g., type of equipment) rather than medical information.[53] Grad-CAM provides the basis of each decision made by deep learning-based decision support tools for tasks such as knee osteoarthritis diagnosis[54] and cervical cancer type classification.[55] This study used Grad-CAM to quantify the contributions of different regions. In the deep learning studies, more attention was given to the areas where organoids were forming and had already formed. It established appropriate trust and confidence in users.

Despite the promising results, this study has several limitations. First, only organoid image data were used for prediction in this study. Some studies have analyzed the clinical and molecular features of patients to identify factors influencing the success rate of constructing organoids.[56,57] Using multimodal data may be a successful strategy for improving accuracy and prediction time because multimodal data provide complementary information.[58] Second, there is a need to address future domain shifts by enriching the diversity of the dataset through collaboration with other institutions and continual learning technology.[59] Third, the interpretability analysis in this study only identified which regions were used for the model's predictions but did not delve into the specific attributes of these regions.

Conflicts of interest

None.

Funding

This work was supported by grants from the National Natural Science Foundation of China (No. 82073129), Chongqing Science and Technology Bureau (Nos. CSTB2023NSCQ-MSX1030, cstc-2021jxjl130032, and cstc2022jxjl20039), Ningxia Key Research and Development Program (No. 2023BEG02037), Talent Program of Chongqing (No. cstc2024rcih-bgzxm0162), Chongqing Health Commission (Nos. 2023ZDXM029 and 2023MSXM043), Scientific and Technological Research Program of Chongqing Municipal Education Commission (No. KJQN20230013), the Project for Enhancing Scientific Research Capabilities of Chongqing University Cancer Hospital (No. 2023nlts005), and Beijing Science & Technology Innovation Fund (No. KC2021-JX-0186-129).

References

1. Drost J, Clevers H. Organoids in cancer research. Nat Rev Cancer 2018;18:407–418. doi: 10.1038/s41568-018-0007-6.
2. Gao D, Vela I, Sboner A, Iaquinta PJ, Karthaus WR, Gopalan A, *et al*. Organoid cultures derived from patients with advanced prostate cancer. Cell 2014;159:176–187. doi: 10.1016/j.cell.2014.08.016.
3. van de Wetering M, Francies HE, Francis JM, Bounova G, Iorio F, Pronk A, *et al*. Prospective derivation of a living organoid biobank of colorectal cancer patients. Cell 2015;161:933–945. doi: 10.1016/j.cell.2015.03.053.
4. Boj SF, Hwang CI, Baker LA, Chio, II, Engle DD, Corbo V, *et al*. Organoid models of human and mouse ductal pancreatic cancer. Cell 2015;160:324–338. doi: 10.1016/j.cell.2014.12.021.
5. Hill SJ, Decker B, Roberts EA, Horowitz NS, Muto MG, Worley MJ, Jr., *et al*. Prediction of DNA repair inhibitor response in short-term patient-derived ovarian cancer organoids. Cancer Discov 2018;8:1404–1421. doi: 10.1158/2159-8290.CD-18-0474.
6. Senkowski W, Gall-Mas L, Falco MM, Li YL, Lavikka K, Kriegbaum MC, *et al*. A platform for efficient establishment and drug-response profiling of high-grade serous ovarian cancer organoids. Dev Cell 2023;58:1106–1121. doi:10.1016/j.devcel.2023.04.012.
7. Bose S, Clevers H, Shen XL. Promises and challenges of organoid-guided precision medicine. Med 2021;2:1011–1026. doi: 10.1016/j.medj.2021.08.005.
8. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–444. doi: 10.1038/nature14539.
9. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, *et al*. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88. doi: 10.1016/j.media.2017.07.005.
10. Kegeles E, Naumov A, Karpulevich EA, Volchkov P, Baranov P. Convolutional neural networks can predict retinal differentiation in retinal organoids. Front Cell Neurosci 2020;14:11. doi: 10.3389/fncel.2020.00171.
11. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, *et al*. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402–2410. doi: 10.1001/jama.2016.17216.
12. Kawahara J, Brown CJ, Miller SP, Booth BG, Chau V, Grunau RE, *et al*. BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. Neuroimage 2017;146:1038–1049. doi: 10.1016/j.neuroimage.2016.09.046.
13. Li JP, Qiu ZX, Cao KY, Deng L, Zhang WJ, Xie CM, *et al*. Predicting muscle invasion in bladder cancer based on MRI: A comparison of radiomics, and single-task and multi-task deep learning. Comput Methods Programs Biomed 2023;233:107466. doi: 10.1016/j.cmpb.2023.107466.
14. Ma MF, Zhang XM, Li Y, Wang X, Zhang RG, Wang Y, *et al*. ConvLSTM coordinated longitudinal transformer under spatio-temporal features for tumor growth prediction. Comput Biol Med 2023;164:107313. doi: 10.1016/j.compbiomed.2023.107313.
15. Yang F, Pham TA, Brandenberg N, Lutolf MP, Ma J, Unser M. Robust phase unwrapping via deep image prior for quantitative phase imaging. IEEE Trans Image Process 2021;30:7025–7037. doi: 10.1109/tip.2021.3099956.
16. Kassis T, Hernandez-Gordillo V, Langer R, Griffith LG. OrgaQuant: Human intestinal organoid localization and quantification using deep convolutional neural networks. Sci Rep 2019;9:12479. doi: 10.1038/s41598-019-48874-y.
17. Bian XS, Li G, Wang C, Liu WQ, Lin XH, Chen ZX, *et al*. A deep learning model for detection and tracking in high-throughput images of organoid. Comput Biol Med 2021;134:11. doi: 10.1016/j.compbiomed.2021.104490.
18. Abdul L, Xu J, Sotra A, Chaudary A, Gao J, Rajasekar S, *et al*. D-CryptO: Deep learning-based analysis of colon organoid morphology from brightfield images. Lab Chip 2022;22:4118–4128. doi: 10.1039/d2lc00596d.
19. Lim MH, Shin S, Park K, Park J, Kim SW, Basurrah MA, *et al*. Deep learning model for predicting airway organoid differentiation. Tissue Eng Regen Med 2023;20:1109–1117. doi: 10.1007/s13770-023-00563-8.
20. Larsen BM, Kannan M, Langer LF, Leibowitz BD, Bentaieb A, Cancino A, *et al*. A pan-cancer organoid platform for precision medicine. Cell Rep 2021;36:109429. doi: 10.1016/j.celrep.2021.109429.
21. Sockell A, Wong W, Longwell S, Vu T, Karlsson K, Mokhtari D, *et al*. A microwell platform for high-throughput longitudinal phenotyping and selective retrieval of organoids. Cell Systems 2023;14:764–766.e6. doi: 10.1016/j.cels.2023.08.002.
22. Simonyan K, Zisserman A, eds. Very deep convolutional networks for large-scale image recognition. 3rd International Conference on

Learning Representations (ICLR 2015); 2015: Computational and Biological Learning Society. San Diego, CA, USA.
23. He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. 2016 IEEE Conf Comp Vis Pattern Recogn 2016;770–778. doi: 10.1109/Cvpr.2016.90.
24. Woo S, Debnath S, Hu R, Chen X, Liu Z, Kweon IS, *et al*, eds. Convnext v2: Co-designing and scaling convnets with masked autoencoders. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2023. San Diego, CA, USA.
25. Liu Z, Hu H, Lin Y, Yao Z, Xie Z, Wei Y, *et al*. Swin Transformer V2: Scaling Up Capacity and Resolution. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)2022. p. 11999–12009. San Diego, CA, USA.
26. Bottou L. Large-Scale Machine Learning with Stochastic Gradient Descent. Compstat'2010: 19th International Conference on Computational Statistics 2010;177–186. doi: 10.1007/978-3-7908-2604-3_16.
27. Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: A literature review. BMC Med Imag 2022;22:69. doi: 10.1186/s12880-022-00793-7.
28. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, *et al*. Attention is all you need. Adv Neural Inf Process Syst 2017;30:5998–6008.
29. Loh HW, Ooi CP, Seoni S, Barua PD, Molinari F, Acharya UR. Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011-2022). Comput Meth Programs Biomed 2022;226:107161. doi: 10.1016/j.cmpb.2022.107161.
30. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. Int J Comput Vis 2019;128:336–359. doi: 10.1007/s11263-019-01228-7.
31. Xiang HL, Xiao YJ, Li F, Li CY, Liu LX, Deng TT, *et al*. Development and validation of an interpretable model integrating multimodal information for improving ovarian cancer diagnosis. Nat Commun 2024;15:2681. doi: 10.1038/s41467-024-46700-2.
32. Bao D, Wang L, Zhou X, Yang S, He K, Xu M. Automated detection and growth tracking of 3D bio-printed organoid clusters using optical coherence tomography with deep convolutional neural networks. Front Bioeng Biotechnol 2023;11:1133090. doi: 10.3389/fbioe.2023.1133090.
33. Okamoto T, Natsume Y, Doi M, Nosato H, Iwaki T, Yamanaka H, *et al*. Integration of human inspection and artificial intelligence-based morphological typing of patient-derived organoids reveals interpatient heterogeneity of colorectal cancer. Cancer Sci 2022;113:2693–2703. doi: 10.1111/cas.15396.
34. Chan CJ, Heisenberg CP, Hiiragi T. Coordination of morphogenesis and cell-fate specification in development. Curr Biol 2017;27:R1024–R1035. doi: 10.1016/j.cub.2017.07.010.
35. Chiaradia I, Imaz-Rosshandler I, Nilges BS, Boulanger J, Pellegrini L, Das R, *et al*. Tissue morphology influences the temporal program of human brain organoid development. Cell Stem Cell 2023;30:1351–1367. doi: 10.1016/j.stem.2023.09.003.
36. Chu SL, Sudo K, Yokota H, Abe K, Nakamura Y, Tsai MD. Human induced pluripotent stem cell formation and morphology prediction during reprogramming with time-lapse bright-field microscopy images using deep learning methods. Comput Methods Programs Biomed 2023;229:107264. doi: 10.1016/j.cmpb.2022.107264.
37. Soelistyo CJ, Vallardi G, Charras G, Lowe AR. Learning biophysical determinants of cell fate with deep neural networks. Nat Mach Intell 2022;4:636–644. doi: 10.1038/s42256-022-00503-6.
38. Liao QY, Zhang Q, Feng X, Huang HB, Xu HH, Tian BY, *et al*. Development of deep learning algorithms for predicting blastocyst formation and quality by time-lapse monitoring. Commun Biol 2021;4:415. doi: 10.1038/s42003-021-01937-1.
39. Lan YQ, Huang NN, Fu YR, Liu KH, Zhang H, Li YZ, *et al*. Morphology-based deep learning approach for predicting osteogenic differentiation. Front Bioeng Biotechnol 2022;9:802794. doi: 10.3389/fbioe.2021.802794.
40. Waisman A, La Greca A, Möbbs AM, Scarafía MA, Velazque NLS, Neiman G, *et al*. Deep learning neural networks highly predict very early onset of pluripotent stem cell differentiation. Stem Cell Rep 2019;12:845–859. doi: 10.1016/j.stemcr.2019.02.004.
41. Zhu YJ, Huang RQ, Wu ZR, Song SM, Cheng LM, Zhu RR. Deep learning-based predictive identification of neural stem cell differentiation. Nat Commun 2021;12:2614. doi: 10.1038/s41467-021-22758-0.
42. D'Sa K, Evans JR, Virdi GS, Vecchi G, Adam A, Bertolli O, *et al*. Prediction of mechanistic subtypes of Parkinson's using patient-derived stem cell models. Nat Mach Intell 2023;5:933–946. doi: 10.1038/s42256-023-00702-9.
43. Rusk N. Deep learning. Nat Meth 2016;13:35-35. doi: 10.1038/nmeth.3707.
44. Mohanty A, Sutherland A, Bezbradica M, Javidnia H. Skin disease analysis with limited data in particular rosacea: A review and recommended framework. IEEE Access 2022;10:39045–39068. doi: 10.1109/Access.2022.3165574.
45. Yue GH, Zhuo GB, Li SY, Zhou TW, Du JF, Yan WQ, *et al*. Benchmarking polyp segmentation methods in narrow-band imaging colonoscopy images. IEEE J Biomed Health Inform 2023;27:3360–3371. doi: 10.1109/Jbhi.2023.3270724.
46. Bansal MA, Sharma DR, Kathuria DM. A systematic review on data scarcity problem in deep learning: Solution and applications. ACM Comput Surv 2022;54:1–29. doi: 10.1145/3502287.
47. Ge Y, Guo YT, Das S, Al-Garadi MA, Sarker A. Few-shot learning for medical text: A review of advances, trends, and opportunities. J Biomed Inform 2023;144:104458. doi: 10.1016/j.jbi.2023.104458.
48. Khan HA, Jue W, Mushtaq M, Mushtaq MU. Brain tumor classification in MRI image using convolutional neural network. Math Biosci Eng 2020;17:6203–6216. doi: 10.3934/mbe.2020328.
49. Al-masni MA, Kim DH, Kim TS. Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. Comp Meth Progr Biomed 2020;190:105351. doi: 10.1016/j.cmpb.2020.105351.
50. Liu Y, Zhang HC, Zhang WD, Lu GJ, Tian Q, Ling N. Few-shot image classification: Current status and research trends. Electronics 2022;11:1752. doi: 10.3390/electronics11111752.
51. Wang M, Deng WH. Deep visual domain adaptation: A survey. Neurocomputing 2018;312:135–153. doi: 10.1016/j.neucom.2018.05.083.
52. Fillioux L, Gontran E, Cartry J, Mathieu JRR, Bedja S, Boilève A, *et al*. Spatio-Temporal Analysis of Patient-Derived Organoid Videos Using Deep Learning for the Prediction of Drug Efficacy. 2023 Ieee/Cvf International Conference on Computer Vision Workshops, Iccvw 2023;3932–3941. doi: 10.1109/Iccvw60793.2023.00425.
53. Barnett AJ, Schwartz FR, Tao CF, Chen CF, Ren YH, Lo JY, *et al*. A case-based interpretable deep learning model for classification of mass lesions in digital mammography. Nat Mach Intell 2021;3:1061–1070. doi: 10.1038/s42256-021-00423-x.
54. Tiulpin A, Thevenot J, Rahtu E, Lehenkari P, Saarakkala S. Automatic knee osteoarthritis diagnosis from plain radiographs: A deep learning-based approach. Sci Rep 2018;8:1727. doi: 10.1038/s41598-018-20132-7.
55. Li YX, Chen F, Shi JJ, Huang YL, Wang M. Convolutional neural networks for classifying cervical cancer types using histological images. J Dig Imaging 2023;36:441–449. doi: 10.1007/s10278-022-00722-8.
56. Xian L, Zhao P, Chen X, Wei Z, Ji H, Zhao J, *et al*. Heterogeneity, inherent and acquired drug resistance in patient-derived organoid models of primary liver cancer. Cell Oncol 2022;45:1019–1036. doi: 10.1007/s13402-022-00707-3.
57. Nuciforo S, Fofana I, Matter MS, Blumer T, Calabrese D, Boldanova T, *et al*. Organoid models of human liver cancers derived from tumor needle biopsies. Cell Rep 2018;24:1363–1376. doi: 10.1016/j.celrep.2018.07.001.
58. Behrad F, Saniee Abadeh M. An overview of deep learning methods for multimodal medical data mining. Expert Syst Appl 2022;200:117006. doi: 10.1016/j.eswa.2022.117006.
59. Perkonigg M, Hofmanninger J, Herold CJ, Brink JA, Pianykh O, Prosch H, *et al*. Dynamic memory to alleviate catastrophic forgetting in continual learning with medical imaging. Nat Commun 2021;12:5678. doi: 10.1038/s41467-021-25858-z.

How to cite this article: Wu HJ, Ma LF, Wang L, Zhu XP, Luo XG, Zhang C, Ha CF, Dang Y, Wang HX, Zou DL. Development and validation of deep learning for predicting the growth of ovarian cancer organoids. Chin Med J 2026;139:108–117. doi: 10.1097/CM9.0000000000003575