# Early prediction of successful organoid establishment using deep learning based microscopic image analysis

**Taeyong Kweon**

**The Graduate School**
**Yonsei University**
**Department of Medical Science**

# Early prediction of successful organoid establishment using deep learning based microscopic image analysis

**A Master's Thesis Submitted**
**to the Department of Medical Science**
**and the Graduate School of Yonsei University**
**in partial fulfillment of the**
**requirements for the degree of**
**Master's of Medical Science**

**Taeyong Kweon**

**June 2024**

**This certifies that the Master's Thesis of Taeyong Kweon is approved.**

Thesis Supervisor Hyung Seok Park

Thesis Committee Member Min Hwan Kim

Thesis Committee Member Nam Kug Kim

**The Graduate School**
**Yonsei University**
**June 2024**

ACKNOWLEDGEMENTS

I'm truly thankful to my family for their constant support and to my research team for their valuable contributions that made our work possible. Special thanks to the AI team for their crucial role in our research. I am grateful to Professor Namkug kim and Min Hwan Kim for their guidance and teaching to proceed this challenging research. I'm especially grateful to Professor Hyung Seok Park for being an exceptional mentor, providing guidance and encouragement.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ABSTRACT

## Early prediction of successful organoid establishment using deep learning based microscopic image analysis

The complexity and heterogeneity of cancer pathogenesis necessitate the development of personalized in-vitro models. Patient-derived organoids (PDOs) offer a promising approach, maintaining patient-specific mutations and enabling individualized therapeutic strategies. However, the establishment success rates of organoids vary significantly across cancer types, posing a challenge especially for large-scale application. This study introduces an advanced AI-based model utilizing deep learning to predict the success of organoid establishment at an early stage using microscopic images. Employing a dual-image input approach combined with Masked Autoencoder (MAE) pretraining, Vision Transformer (ViT) models demonstrated high predictive accuracy, surpassing human performance in several metrics. The ViT-Dual-MAE-Finetune and Linearprobe models were particularly effective, achieving the highest Area Under the Curve (AUC) scores of 0.88 with total test set. The ViT-Dual-MAE-Linearprobe model consistently improved across organoid passages, reaching a peak AUC of 0.98 in Passage 2. Furthermore, the models performed well with HER2 and TNBC subtypes. Attention map analysis revealed that successful predictions often focused on the edges of individual organoids, suggesting that these features may be critical indicators of growth. These findings highlight the potential of AI models in enhancing the efficiency and scalability of organoid-based research, paving the way for significant advancements in precision medicine and personalized cancer treatment.

---

Key words : patient derived organoids; artificial intelligence; organoid establishment success rate;

# 1. INTRODUCTION

As oncological research and genomic technology advances, the high complexity and heterogeneity of cancer pathogenesis has been revealed[1]. It is now understood that what was previously categorized as a single type of cancer is actually a collection of various distinct subgroups[2]. Even patients with the same subtype of breast cancer may show different treatment responses due to unique genetic mutations[3]. Considering this characteristic of cancer, the development of an in-vitro model for personalized cancer treatment has become essential[4].

PDO which are three-dimensional cell cultures mimicking the architectural and genetic landscape of the originating tumors, can fulfill this need by maintaining the patient-specific mutations and expression profiles[5]. This ability to recapitulate the tumor characteristics enables individualized therapeutic strategies for each cancer profile[6]. Therefore, the development of organoid-based disease models is gaining traction as a platform for future precision medicine[7].

The PDO disease model Platform, which is often called a living biobank in oncological research, goes beyond the establishment of individual organoids[8,9]. It involves creating a comprehensive set of PDOs that collectively represent a diverse range of patient-specific tumor mutational variations. This extensive collection of organoids not only enables precision medicine through personalized drug response experiments but also serves as a predictive tool in the preclinical stage of new drug development. It can significantly reduce clinical trial costs and enhance efficiency, thereby promote the path to innovative cancer treatments[10,11].

The establishment of PDO marks a significant advancement, offering considerable potential for drug discovery, disease modeling, and personalized medicine. However, a critical challenge that persists across this domain, the establishment success rate of tumor organoids across various cancer types has not achieved the 100%: prostate cancer organoids exhibit a success rate of approximately 20%[12], breast cancer organoids range from 77% to 87.5%[13,14], pancreatic cancer organoids have a success rate of about 75% to 83%[15], colorectal and lung cancer organoids show success rates of 63%[16] and 75%[17] respectively. These figures highlight the significant impact of these success rates on the scalability of organoid establishment techniques at a platform level.

A key method to improve the efficiency of organoid establishment is to identify morphological characteristics that distinguish successful organoids from unsuccessful ones[18]. However, manual identification of these characteristics by human observers presents significant obstacles. Organoids

may be located closely due to high-density which makes it difficult to distinguish one another. The 3D cultivation of organoids leads to a distribution along the z-axis, which complicates the capture of all organoid information on a single focal plane. Furthermore, the heterogeneity in organoid size and shape, which is due to diverse proliferation rates of organoid cells[19]. This is where the role of Artificial intelligence (AI) becomes indispensable.

Recently, several pioneering platforms have applied AI technologies, particularly machine learning and deep learning, to organoid research. OrganoID is an image analysis platform that automatically recognizes, labels, and tracks single organoids in microscopy images[20]. OrgaTracker utilizes semantic segmentation to track and segment organoids over time, capturing and analyzing organoid fusion. D-CryptO is designed for organoid morphology classification, assessing crypt formation and the opacity of colorectal organoids[21]. OrganoSeg combines advanced segmentation, filtering, and analysis tools tailored for archived 3D cultures of organoid [22].

While these AI approaches to organoid image primarily focus on academic data analysis, including morphological analysis and segmentation of single organoids, their contribution to the practical efficiency of experimental operations has been limited[20,21,23]. Our research introduces a novel AI model designed to predict the success or failure of organoid establishment at an early stage. By utilizing simple microscopy images at two different time points, this two ViT fused architectural AI model enables researchers to find out the likelihood of an organoid's successful establishment in the early stage. By identifying in early stage of organoid establishment whether an organoid is likely to be successful or fail, resources can be efficiently distributed to the cases with a higher probability of success. Furthermore, by analyzing the morphological characteristics of cases predicted to be successful or failed by a deep learning model, we can obtain valuable biological insights, enabling a deeper understanding. This predictive capability is a significant step towards optimizing organoid culture processes and enhancing the practicality of organoid-based research.

# 2. MATERIALS AND METHODS

Figure 1 illustrates the comprehensive workflow from the initial culturing of organoids to the preparation of the dataset for deep learning analysis.

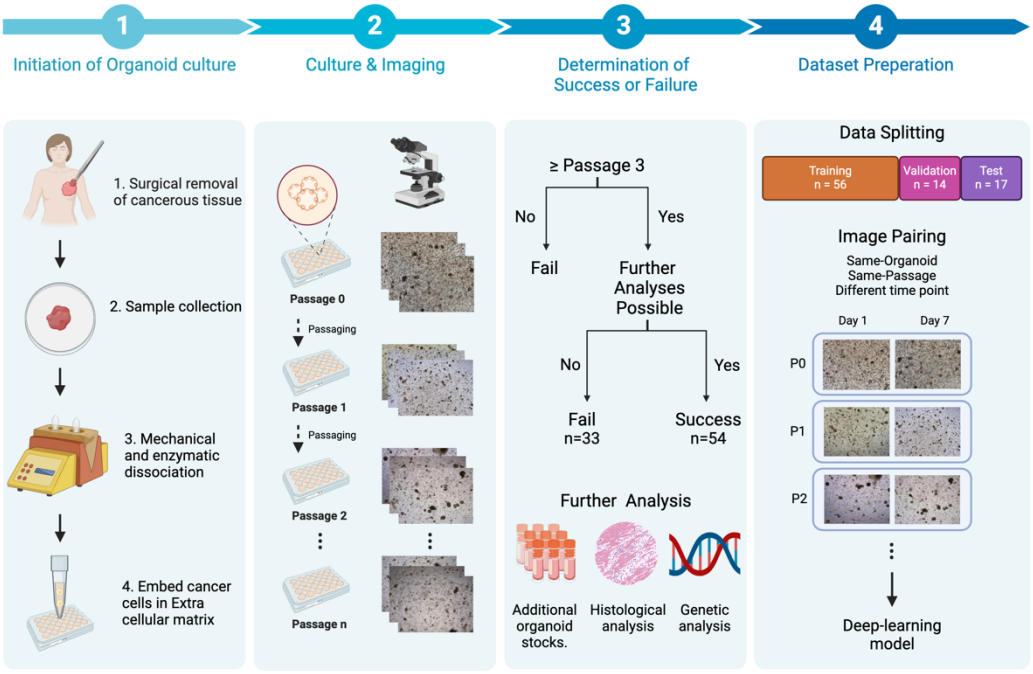



**Figure 1. Workflow from organoid culture to dataset preparation for deep learning model.** Cancer tissues are surgically extracted, dissociated mechanically and enzymatically, and embedded in an extracellular matrix to grow organoids. These organoids are passaged and imaged at each stage. Organoid establishment is classified as successful or failed based on predefined criteria. A total of 87 organoids are categorized into training, validation, and test sets, with images paired according to specific rules to create the dataset for deep learning analysis.

## 2.1. Breast cancer organoid culture

The Cancer tissue samples were obtained through surgical resection. mechanical mincing of the tissue was performed by using a scalpel blade on a Petri dish. For chemical dissociation, minced tissue fragments were transferred into a MACS tube (Miltenyi, Auburn, CA, USA) with 1ml of Collagenase (Sigma-Aldrich, St. Louis, MO, USA), 2μl of Y-27632 (Selleckchem, Houston, TX, USA), and 3ml of AdDf+++ (GIBCO, Grand Island, NY, USA). The samples were then dissociated for 1 hour using a gentleMACS Octo Dissociator (Miltenyi Biotec, Bergisch Gladbach, Germany) with Heater. The solution in the MACS tube was filtered through a 100μm pore size strainer. The filtrate was supplemented with 5ml of ADDF and 200μl of FBS (Thermo Fisher Scientific, Waltham, MA, USA), and then centrifuged at 400rcf for 3 minutes at 25°C. 2ml of RBC lysis (Roche, Basel, Switzerland) buffer was applied for 5 minutes at room temperature, followed by washing process with 2ml of PBS (Thermo Fisher Scientific, Waltham, MA, USA). The washed pellet was then mixed with BME2 (Cultrex, R&D Systems, Minneapolis, MN, USA) and AdDf+++ in a 9:1 ratio and seeded into a 24-well plate (Thermo Fisher Scientific, Waltham, MA, USA).

## 2.2. Organoid media

The organoid culture medium was prepared as follows: To 500 ml of Advanced DMEM/F12, 5 ml of penicillin–streptomycin (Thermo Fisher Scientific, Waltham, MA, USA), 5 ml of GlutaMAX (Thermo Fisher Scientific, 1% vol/vol), and 5 ml of HEPES (GIBCO, Grand Island, NY, USA) were added for basic supplementation. Further additions to the medium included 30% v/v L-WRN conditioned medium (ATCC, cat. no. CRL-3276), Primocin (InvivoGen, San Diego, CA, USA), nicotinamide (Sigma-Aldrich, St. Louis, MO, USA), N-acetylcysteine (Sigma-Aldrich, St. Louis, MO, USA), and hydrocortisone (Sigma-Aldrich, St. Louis, MO, USA). Additional components included A83-01 (Tocris, Minneapolis, MN, USA), Y-27632, forskolin (TCI, Tokyo, Japan), β-estradiol (Sigma-Aldrich, St. Louis, MO, USA), heregulin β1 (PeproTech, Rocky Hill, NJ, USA), FGF 10 (PeproTech, Rocky Hill, NJ, USA), EGF (PeproTech, Rocky Hill, NJ, USA), and 1x B27 supplement (Thermo Fisher Scientific, Waltham, MA, USA).

## 2.3. Organoid passaging

Media change was conducted every 3 to 4 days and organoid was subcultured depending on growth rate of each organoid at a ratio of 1:2 to 1:3. The process of passaging involves collecting organoids from each well with 400ul of PBS. Harvesting solution (Cultrex, R&D Systems, Minneapolis, MN, USA) was applied to the mixture of organoids and Matrigel for 1hour at an ice-cold temperature with 110 rpm of laboratory shaker. Organoids were then washed with 2ml of PBS, dissociated into appropriate sizes using Tryple Express (Thermo Fisher Scientific, Waltham, MA, USA), followed by PBS wash. Organoids were seeded in 24well plate with a BME2: AdDF+++ mixture at a 9:1 ratio.

## 2.4. Image acquisition

For the acquisition of images, organoids seeded in 24-well plates and treated with media were visualized using an Olympus IX73 inverted microscope. Image capture was facilitated by using the CellSens Standard software. The imaging conditions included the use of a 4x objective lens combined with a 10x ocular lens under optical conditions. The acquired images were recorded at a resolution of 1600x1200 pixels and saved in JPG format.

## 2.5. Datasets

Our research utilized a comprehensive dataset comprising a total of 29,847 images, which were derived from 87 distinct organoids. Each organoid originated from unique patient breast cancer tissues, ensuring a wide representation of the disease's variability. Among these, two cases were harvested using a core needle biopsy technique. The remaining organoids were obtained from tissues excised during surgical procedures. This dataset encompassed 54 cases that were classified as successful organoid cultures and 33 that were deemed unsuccessful.

## 2.6. Establishment criteria

The success of organoid establishment was set under two conditions. First, organoid should reach passage 3 which indicate that the organoid successfully undergone three subculture process and feasible of stable long-term culture. Second, considered to be possible for further subculture and histological, genetic evaluations.

## 2.7. Image pair

Our ViT-Dual models use pairs of images as input, to improve model performance. These pairs were randomly selected from the same organoid and passage, with a temporal difference of 3 to 10 days between them. Any duplicate image pairs were removed before training and testing.

## 2.8. Train, Test and Validation set

The organoid dataset was divided into training, validation, and test sets, consisting of 56, 14, and 17 samples each, respectively. The training subset contained 19910 image pairs, 5012 pairs for the validation subset and 3941 pairs for the test subset. The discrepancy in the ratio of the number of samples to the number of images is due to the variable duration required for the establishment of each sample, which affects the quantity of images captured as each organoid grows.

## 2.9. Preprocessing and augmentation

The images undergo a series of preprocessing steps before being input into the model, including resizing to 512x512 pixels, normalization, and augmentations (random horizontal flips, vertical flips, and rotations for training data).

## 2.10. Train, Test and Validation set

The organoid dataset was divided into training, validation, and test sets, consisting of 56, 14, and 17 samples each, respectively. The training subset contained 19910 image pairs, 5012 pairs for the validation subset and 3941 pairs for the test subset. The discrepancy in the ratio of the number of samples to the number of images is due to the variable duration required for the establishment of each sample, which affects the quantity of images captured as each organoid grows.

## 2.11. Training configuration

The main models were initially pretrained using a MAE. This pretraining approach involves randomly masking 75% of the input image data and training the model to reconstruct the original unmasked image. After pretraining, the weights of the ViT were fixed. In the subsequent main

training phase, only the classifying part of the model were trained. The model was trained for 100 epochs with a batch size of 128. The Adam optimizer was used with a learning rate of 1e-4. CrossEntropyLoss was employed as the loss function to guide the training process. The model's performance was validated with a separate validation set, and the best-performing model was saved for further evaluation on the test set.

## 2.12. Control model using single image input

The control model utilizes the same ViT architecture but is processing single images to predict organoid establishment success. The control model's training, validation and test set included same patients with the main model's dataset. The model was trained for 25 epochs with a batch size of 32 to see the training aspect. Adam optimizer with a learning rate of 1e-4 and crossentropyloss as the loss criterion. The performance of the control model and the main model was compared and evaluated based on AUC values.

## 2.13. Human prediction

The human prediction analysis was conducted on a subset of image pairs from the test set that corresponded to Passage 0,1 and 2. This task was carried out by two human breast cancer researchers. To mitigate potential bias that could arise from predicting images of a specific passage in sequence, the order of image pair was randomly shuffled, and any passage and day information was given to researcher.

## 2.14. Statistical analysis

For statistical analyses, different tests were employed based on the nature of the data and the specific comparisons being made:

AUC Comparison: DeLong's test was used for comparing the AUCs between models on the same samples. For AUC comparisons with different samples, a permutation test was applied. Accuracy Comparison: McNemar's test was utilized to compare the accuracy between models. Precision and Recall: The Wilcoxon Signed Rank test was conducted to compare precision and recall between models. Duration Analysis: An ANOVA test was performed to analyze the average

differences in duration between groups. This was followed by the Tukey HSD test to identify statistically significant differences.

A p-value of less than 0.05 was considered statistically significant for all tests. All statistical analyses were performed using Python 3.11.2, specifically employing the following libraries: scipy.stats for the ANOVA test and statsmodels.stats.multicomp for the Tukey HSD test, sklearn.metrics for accuracy, precision, and recall calculations, statsmodels.stats.contingency_tables for McNemar's test, and scipy.stats for the Wilcoxon Signed Rank test and permutation test.

## 2.15. Attention map

To understand the decision-making process of our Vision Transformer model, we visualized the attention map from the final layer, highlighting the most important image regions influencing the model's classification by overlaying attention scores as a heatmap. To facilitate comparison of attention maps, a graphical user interface (GUI) was developed using the 'Tkinter' library in Python. This interface allowed users to load pairs of images and their corresponding attention maps. The GUI was made to adjust the transparency of the attention map overlays.

## 2.16. Illustration tool

Figure 1, Figure 3A, and Figure 4B were illustrated using BioRender (BioRender.com).

## 2.17. Code and software

The codes generated and/or analyzed during the current study are available in the GitHub repository: https://github.com/ghdwnsgh852/orgapredictor.

## 2.18. Data avilability

Due to privacy concerns, the dataset is not publicly available. But it can be shared upon request for research purposes. Researchers interested in accessing the dataset can submit a request to Taeyong Kweon at kmskty3@yuhs.ac.

# 3. Results

## 3.1. Tracking organoid development

The initiation point of organoid culture, including subculturing for passage progression, was designated as Day 0. Imaging conducted approximately every 7 days enabled the tracking of morphological evolution over time. Consistent imaging locations within each passage facilitated accurate monitoring of organoid development (Figure 2). Observations revealed that organoids increased in size and cell density within each passage, helping identify appropriate passaging times, monitor for contamination, and recognize growth failures.

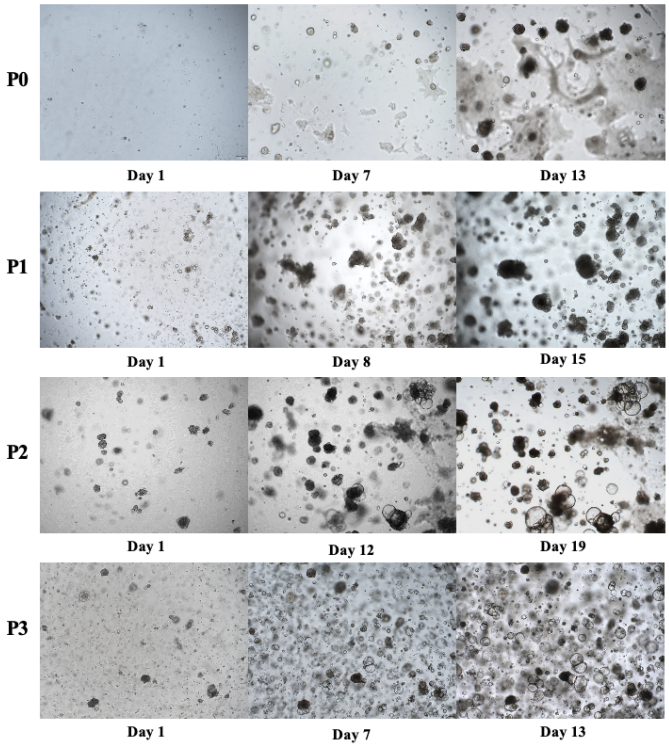



**Figure 2. Organoid culture progression and microscopy images.** The displayed images illustrate the progression of organoid culture from Passage 0 to Passage 3 in a case of Luminal A breast cancer with successful establishment, captured approximately every 7 days at the discretion of the researcher. Each image represents a representative image selected from multiple positions and focal planes.

## 3.2. Imaging for comprehensive organoid analysis

To account for the inherent morphological diversity, multiple images were captured from at least two distinct positions for each organoid (Figure 3A). Additionally, considering the three-dimensional architecture, imaging was conducted at various focal planes along the z-axis. This approach ensured a comprehensive capture of organoids from different positions and depths (Figure 3B).

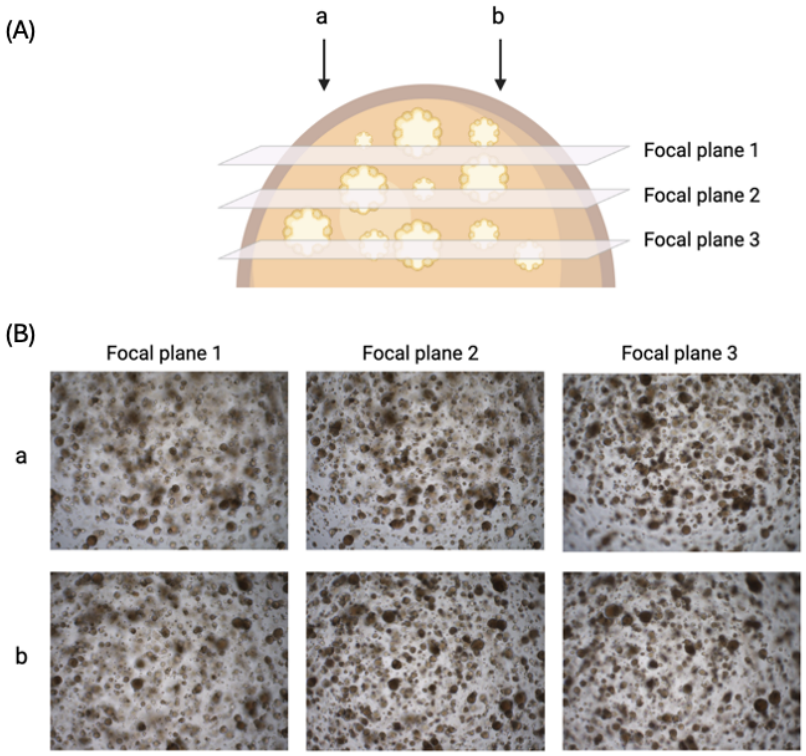



**Figure 3. Organoid imaging through multiple positions and focal planes.** (A) Schematic representation of an organoid: Illustrates a hemispherical organoid possessing a 3D structure along the z-axis. Position a and b indicate multiple imaging locations on a single organoid. The focal plane is depicted as planes at various depths along the z-axis. (B) Microscopy images according to focal planes: Displays the microscopy images for each of the positions a and b, corresponding to the focal planes at different depths along the z-axis.

### 3.3. Characteristics of study population and Image Dataset

Between May 2021 and February 2024, we obtained a large dataset of 36,824 breast cancer organoid microscopy images in the process of establishing organoids. These images were collected to monitor the organoids' growth, determine their viability, check for contamination, and identify the appropriate time points for subculture. Each image was labeled with patient ID, passage number, and the day of capture to facilitate tracking. The images were annotated based on pre-established criteria for successful organoid establishment.

Initially, 36,824 images were collected, but 6,977 images were excluded due to issues such as ongoing construction at the time of analysis, bubbles in the gel, different magnifications, excessive inclusion of non-gel areas, contamination, and images taken after re-culturing from stock. Breast cancer organoid microscopy images were collected from a single patient population (N = 87) at Yonsei Severance Hospital. Out of the 87 organoids, 54 cases successfully established organoids, while 33 cases were unsuccessful (Table 1).

For data analysis, we divided the dataset into training, testing, and validation sets based on patients, not images, to ensure that all images from the same patient were in the same set. This approach was chosen because images from the same passage were captured over time and thus should not be split across different sets. Each image was paired to serve as dual input for a dual vision transformer model to achieve higher predictive performance compared to single input. These pairs were created based on specific criteria (see "Methods" section 2.7). The training set includes 19,910 image pairs from 56 patients, the test set includes 5,012 image pairs from 14 patients, and the validation set includes 3,941 image pairs from 17 patients. The breast cancer subtypes were considered, with 17 cases of Luminal A, 31 cases of Luminal B, 9 cases of HER2-positive, and 30 cases of triple-negative included in the analysis.

**Table 1. Characteristics of study population and Image Dataset**

| Characteristic | Training Set | Testing Set | Validation Set | Total |
|---|---|---|---|---|
| Patients | 56 | 17 | 14 | 87 |
| Image pairs | 19,566 | 3,941 | 4,928 | 28,435 |
| Success Cases | 34 | 12 | 10 | 56 |
| Failure Cases | 22 | 5 | 4 | 31 |
| Luminal A | 9 | 5 | 3 | 17 |
| Luminal B | 22 | 4 | 5 | 31 |
| HER2 | 5 | 3 | 1 | 9 |
| Triple-negative | 20 | 5 | 5 | 30 |

### 3.4. Model architecture using Vision Transformer for organoid analysis

In this study, we employed a ViT model to predict the success of organoid establishment. The ViT was chosen for its ability to generate attention maps, which highlight the regions of the image that the model focuses on during prediction[24]. This feature provides both predictive capabilities and biological insights into organoid development.

Unlike traditional models that use a single image input, our approach uses a dual-image input. This method captures subtle morphological features in organoids, to improve the model's predictive accuracy. Images from the same organoid and passage, taken at different time points (3-10 days apart), were paired and fed into a two ViT architecture. The output features from the transformer were concatenated and passed through a fully connected layer to produce a binary prediction of success or failure (Figure 4A).

To enhance the model's ability to detect features, the ViT backbone was pretrained using a MAE. During this pretraining phase, up to 75% of the original image data was masked, and the model was trained to reconstruct the obscured portions (Figure 4B). This process equipped the model with a robust representation of organoid features, crucial for recognizing key morphological indicators[25].

After MAE pretraining, the model underwent two different training methodologies: fine-tuning and linear probing. In fine-tuning, all parameters, including those of the multilayer perceptron (MLP) components, were trained. This comprehensive approach allowed the model to adjust its weights based on the pretrained features. In linear probing, only the MLP layer was trained while the transformer weights remained fixed, maintaining the pretrained feature representations.

The ViT model adapts transformer mechanisms, traditionally used for sequential data, to image analysis. Each image was divided into a grid of 16x16 pixel patches, which were then flattened, linearly projected, and augmented with positional embeddings to preserve spatial context (Figure 4C). The transformer encoder processed these inputs through multi-headed self-attention mechanisms, enabling the model to focus on multiple segments of the image simultaneously. Normalization layers stabilized the learning process, and MLP layers integrated the processed features. This architecture allowed the model to detect subtle morphological features in organoids, recognizing patterns and changes over time to predict successful establishment.

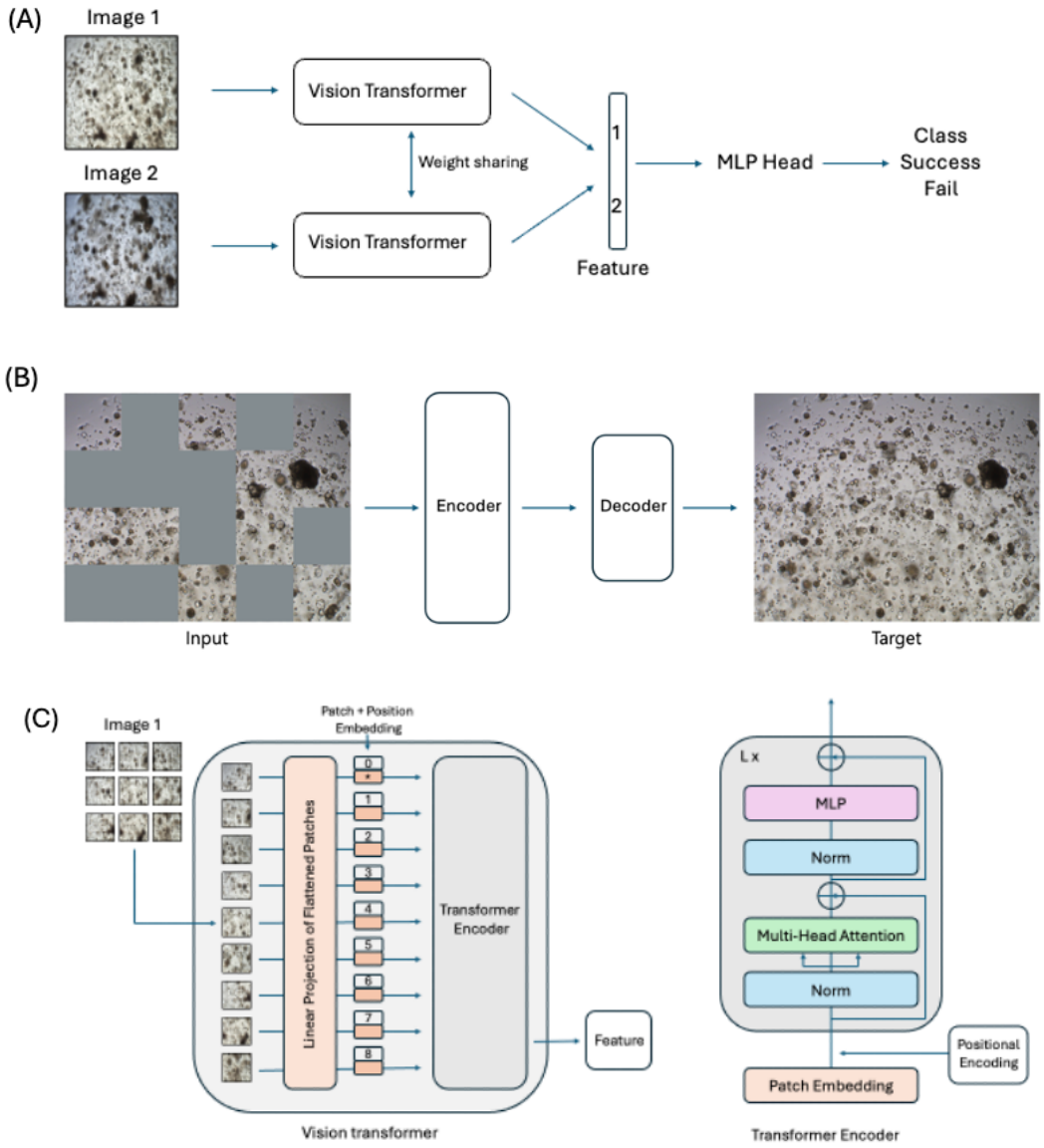



**Figure 4. Vision Transformer model architecture for dual image input and Masked Autoencoder.** (A) Schematic of the ViT-Dual-MAE model structure used in this study: it illustrates how two input images are fed into a weight-sharing ViT to generate features. These features are then concatenated and used to predict two classes. (B) The structure of the MAE: This schematic shows

the MAE's process where the input image is partially masked, then passed through the encoder to generate a latent representation. The decoder reconstructs the image from this representation, aiming to predict the original unmasked image. (C) Schematic of the ViT and Transformer Encoder structures. *Norm; normalization, MLP; multi-layer perceptron.

### 3.5. Comparative analysis of Vision Transformer Models

Four ViT models were evaluated to determine the optimal approach for predicting organoid establishment success: ViT-Single, ViT-Dual, ViT-Dual-MAE-Finetune, and ViT-Dual-MAE-Linearprobe Model. The performance of these models was assessed using the Area Under the Receiver Operating Characteristic Curve (AUC), and the statistical significance of their differences was evaluated using appropriate tests.

The ViT-Single Model, which processes a single image input, achieved an AUC of 0.72 (95% CI 0.70-0.74) on the test set (Figure 5). This provided a baseline for assessing the benefits of incorporating dual-image inputs and MAE pretraining.

The ViT-Dual Model, using dual image inputs without MAE pretraining, showed improved performance with a Test AUC of 0.84 (95% CI 0.82-0.86), which significantly improved the model performance (permutation test, $p < 0.001$).

Further advancements were observed with the MAE-pretrained models. The ViT-Dual-MAE-Finetune Model achieved a Test AUC of 0.88 (95% CI 0.86-0.89), and the ViT-Dual-MAE-Linearprobe Model achieved a Test AUC of 0.88 (95% CI 0.86-0.90), both of which had significantly higher AUCs than the ViT-Dual Model without MAE pretraining (DeLong's test, both $p < 0.001$).

This suggests that MAE pretraining enhances the model's capability to extract relevant features from microscopy images, contributing to more accurate predictions.

Overall, the results demonstrate that the ViT-Dual models, particularly those with MAE pretraining, provide superior predictive performance. The dual-image input approach, combined with the MAE pretraining, appears to improve the capture of critical morphological features necessary for predicting organoid establishment success.

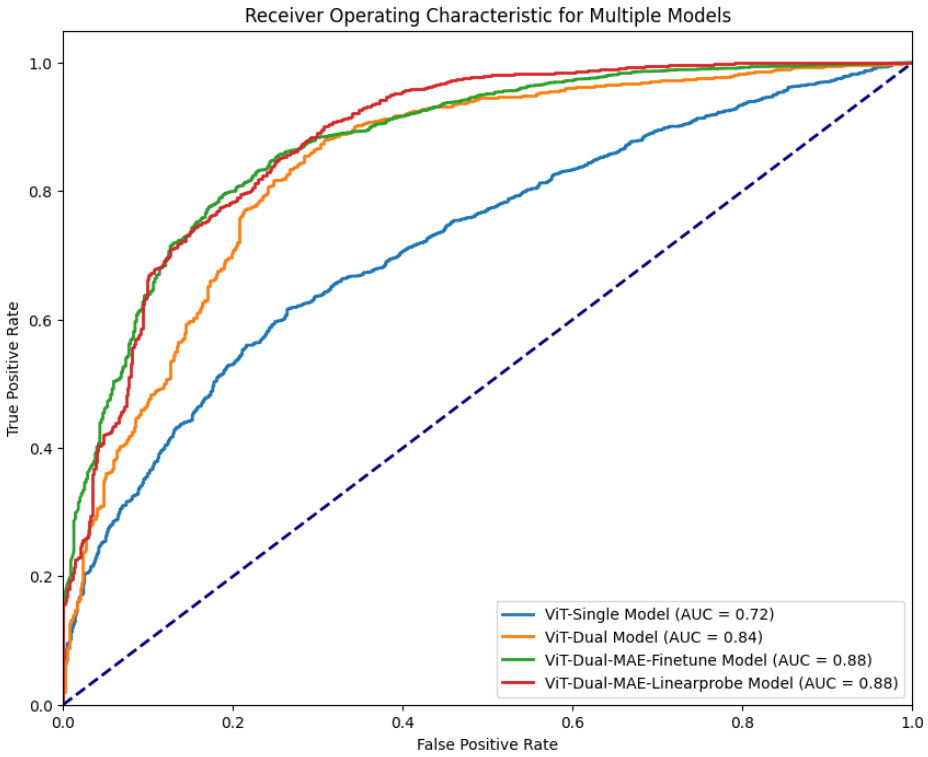



**Figure 5. Comparative ROC Curves of Vision Transformer Models.** Receiver Receiver Operating Characteristics (ROC) curves evaluating the performance of four ViT models in predicting organoid establishment success. The ViT-Single Model (Blue line) achieved an AUC of 0.72 (95% CI 0.70-0.74). The ViT-Dual Model (Orange line) showed an improved AUC of 0.84 (95% CI 0.82-0.86). Both the ViT-Dual-MAE-Finetune Model (Green line) and the ViT-Dual-MAE-Linearprobe Model (Red line) achieved the highest AUC of 0.88, with CIs of 95% (0.86-0.89) and 95% (0.86-0.90) respectively.

### 3.6. Evaluation of ViT-Dual-MAE Models across organoid passages

The performance of the ViT-Dual-MAE-Finetune Model and the ViT-Dual-MAE-Linearprobe Model was further evaluated across different organoid passages to assess their predictive effectiveness at various stages of organoid establishment. The test dataset was segmented into passages 0, 1, and 2, and the AUC was calculated for each passage.

For Passage 0, the ViT-Dual-MAE-Finetune Model exhibited a higher AUC of 0.87 (95% CI: 0.84-0.90) compared to the ViT-Dual-MAE-Linearprobe Model, which achieved an AUC of 0.74 (95% CI: 0.70-0.78; DeLong test, $p < 0.001$) (Figure 6A). This indicates that the fine-tune model is more effective at predicting organoid outcomes at the very early stage of establishment.

In Passage 1, the ViT-Dual-MAE-Linearprobe Model outperformed the fine-tune model, achieving a higher AUC of 0.93 (95% CI: 0.91-0.95) compared to the fine-tune model's AUC of 0.82 (95% CI: 0.78-0.86; DeLong test, $p < 0.001$) (Figure 6B). This suggests that the Linearprobe model's approach of utilizing fixed pretrained transformer weights is more effective at this stage.

For Passage 2, the ViT-Dual-MAE-Linearprobe Model continued to demonstrate superior performance with an AUC of 0.98 (95% CI: 0.97-0.99), while the ViT-Dual-MAE-Finetune Model also showed strong predictive capability with an AUC of 0.94 (95% CI: 0.93-0.96; DeLong test, $p < 0.001$) (Figure 6C). These results highlight the differing effectiveness of the two models at various stages of organoid development.

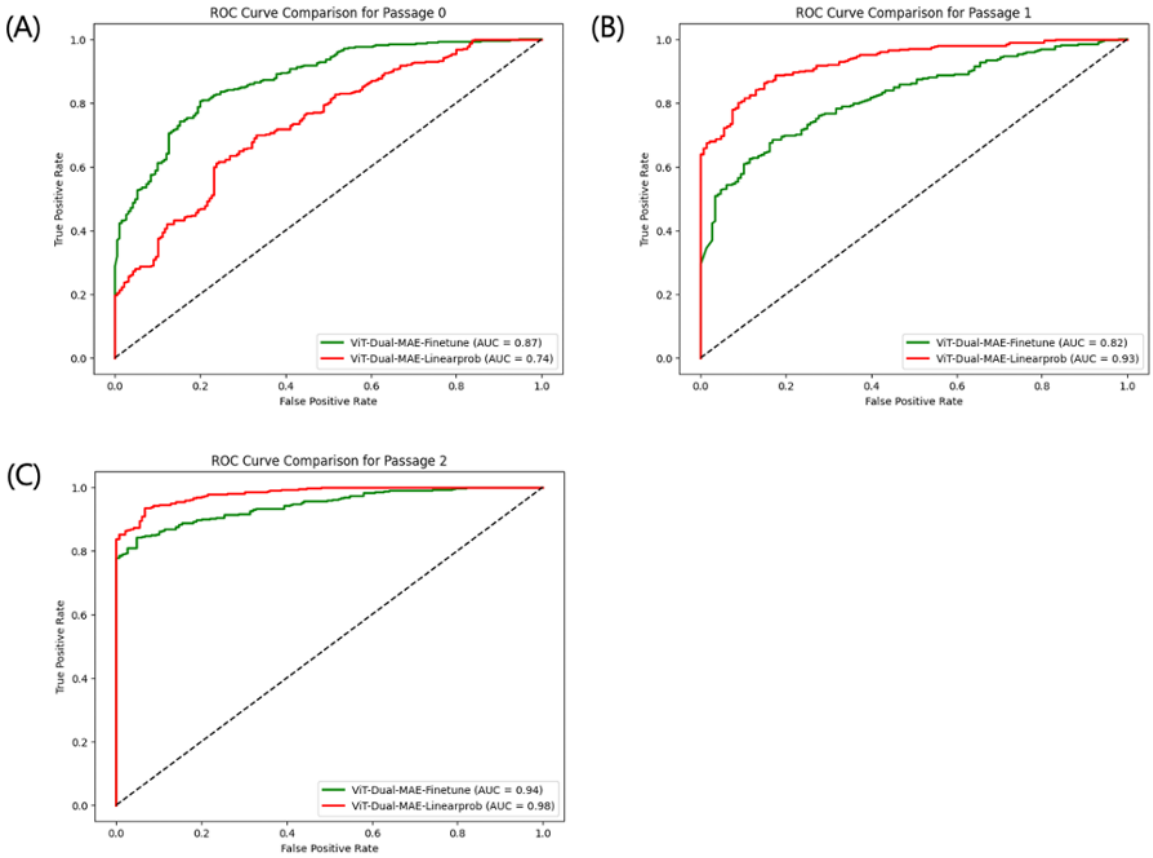



**Figure 6. ROC curves across different organoid passages for ViT-Dual-MAE models.** ROC curves evaluating the performance of two ViT-Dual-MAE models across various passages of organoid establishment. (A) For Passage 0, the ViT-Dual-MAE-Finetune Model (Green line) achieved an AUC of 0.87, while the ViT-Dual-MAE-Linearprobe Model (Red line) achieved an AUC of 0.73. (B) For Passage 1, the ViT-Dual-MAE-Linearprobe Model outperformed with an AUC of 0.93 compared to the ViT-Dual-MAE-Finetune Model's AUC of 0.82. (C) In Passage 2, the ViT-Dual-MAE-Linearprobe Model showed an AUC of 0.98, whereas the ViT-Dual-MAE-Finetune Model recorded an AUC of 0.94.

### 3.7. Comparative performance analysis: AI models vs Human predictions

The performance of two AI models, ViT-Dual-MAE-Finetune and ViT-Dual-MAE-Linearprobe, was compared with that of two organoid experts across three organoid passages (Passage 0, 1, and 2). The evaluation focused on accuracy, precision, recall, F1 score, and AUC metrics (Table 2).

In Passage 0, Human 1 achieved a significantly higher accuracy of 78.49%, while Human 2 had a lower accuracy of 71.60% (McNemar's Test, $p < 0.001$). This indicates a notable difference in the ability to predict success and failure in organoid cultures even among experienced researchers. Similarly, both the ViT-Dual-MAE-Finetune and ViT-Dual-MAE-Linearprobe models demonstrated significantly higher accuracy than Human 2 (McNemar's Test, $p < 0.001$). Human 2 demonstrated the highest precision at 86.56%, indicating fewer false positives; however, this came at the cost of a lower recall rate of 68.04%, suggesting more false negatives (Wilcoxon signed-rank Test, $p < 0.001$). Overall, the ViT-Dual-MAE-Finetune model achieved the highest F1 score, reflecting a balanced performance across precision and recall (Wilcoxon signed-rank Test, $p < 0.001$).

In Passage 1, the ViT-Dual-MAE-Linearprobe model showed improvement with the highest accuracy of 84.73%. Human 2 also improved significantly, achieving an accuracy of 82.74%, which was much higher than that of Human 1 and the ViT-Dual-MAE-Finetune model (McNemar's Test, $p < 0.001$). The Linearprobe model also had the highest recall of 93.06%, along with a commendable precision, resulting in a high F1 score (Wilcoxon signed-rank Test, $p < 0.001$). This indicates effective identification of true positives with minimal false negatives. However, the performance of the ViT-Dual-MAE-Finetune model declined in this passage, indicating challenges.

By Passage 2, the ViT-Dual-MAE-Linearprobe model reached peak performance with perfect precision, achieving no false positives, and an impressive AUC of 98.38. Human 2 had high accuracy at 93.74%, with an exceptional F1 score of 95.66%, reflecting a strong balance between precision and recall. The ViT-Dual-MAE-Finetune model improved from Passage 1 but still lagged behind the ViT-Dual-MAE-Linearprobe model, with an AUC of 94.42.

Based on the results, the ViT-Dual-MAE models, particularly the Linearprobe model, generally outperform human predictors. Notably, in Passage 2, the ViT-Dual-MAE-Linearprobe model achieved perfect precision and a high AUC score, surpassing human performance. These results suggest that AI models have the potential to effectively support human researchers.

**Table 2. AI vs. Human performance in early organoid passages**

| Passage | Metric | Human | | ViT-Dual-MAE | |
|---|---|---|---|---|---|
| | | 1 | 2 | Linearprobe | Finetune |
| | | (%) | (%) | (%) | (%) |
| 0 | Accuracy | 78.49 | 71.60 | 76.68 | 78.98 |
| | Precision | 83.80 | 86.56 | 83.33 | 82.21 |
| | Recall | 84.02 | 68.04 | 81.12 | 87.23 |
| | F1 Score | 83.91 | 76.19 | 82.21 | 84.65 |
| | AUC | - | - | 78.92 | 87.44 |
| 1 | Accuracy | 73.68 | 82.74 | 84.73 | 73.55 |
| | Precision | 76.33 | 86.19 | 85.76 | 75.26 |
| | Recall | 88.85 | 88.85 | 93.06 | 91.17 |
| | F1 Score | 82.12 | 87.50 | 89.26 | 82.45 |
| | AUC | - | - | 93.96 | 82.04 |
| 2 | Accuracy | 86.04 | 93.74 | 93.09 | 87.50 |
| | Precision | 90.81 | 100.00 | 100.00 | 94.63 |
| | Recall | 90.62 | 91.68 | 90.83 | 88.43 |
| | F1 Score | 90.72 | 95.66 | 95.19 | 91.42 |
| | AUC | - | - | 98.38 | 94.42 |

*AUC; area under the curve, AI; artificial intelligence.

## 3.8. Analysis of model performance by breast cancer subtype

The performance of the ViT-Dual-MAE-Finetune and ViT-Dual-MAE-Linearprobe Models was evaluated across four main breast cancer subtypes: Luminal A, Luminal B, Human Epidermal Growth Factor Receptor 2 Positive (HER2), and Triple Negative Breast Cancer (TNBC). The evaluation focused on accuracy, precision, recall, and F1 scores (Table 3).

For the HER2 subtype, both the ViT-Dual-MAE-Finetune and ViT-Dual-MAE-Linearprobe models showed the highest performance above all subtypes (McNemar's Test, $p<0.01$), both achieving an accuracy of 90.27%. Other metrics in this subtype showed similar trends between the two models, with both models performing strongly across precision and recall metrics.

In the Luminal A subtype, the ViT-Dual-MAE-Linearprobe Model exhibited stronger predictive capabilities, registering higher accuracy of 89.35% (McNemar's Test, $p<0.01$) and a high precision of 97.70%, resulting in an F1 score of 93.66% (Wilcoxon Signed Rank Test, $p<0.01$). Conversely, the ViT-Dual-MAE-Finetune Model showed reduced effectiveness, with an accuracy of 75.46% and a recall of 79.10%, highlighting its limitations in accurately identifying true positives compared to the Linearprobe model.

For the TNBC subtype, the ViT-Dual-MAE-Linearprobe Model maintained high accuracy at 83.10% and a high precision of 98.56%. The ViT-Dual-MAE-Finetune Model showed better performance in terms of accuracy to 85.77% and demonstrated balanced high performance in other metrics, all above 90% (McNemar's Test, $p<0.01$; Wilcoxon Signed Rank Test for precision and recall, $p<0.01$).

The Luminal B subtype presented challenges for both models. The ViT-Dual-MAE-Linearprobe Model achieved the lowest accuracy at 76.81%, with a precision of 54.80%, but managed a recall of 100.00%, indicating its ability to identify all true positives despite a high rate of false positives, reflected in a lower F1 score. The performance decline was more pronounced in the ViT-Dual-MAE-Finetune Model, which showed an accuracy of 69.28%, precision of 47.72%, and an F1 score of 63.00%, the lowest across all subtypes. This pattern suggests a high rate of false positives, where the models incorrectly predicted unsuccessful organoid cases as successful (McNemar's Test, $p<0.01$; Wilcoxon Signed Rank Test, $p=0.014$).

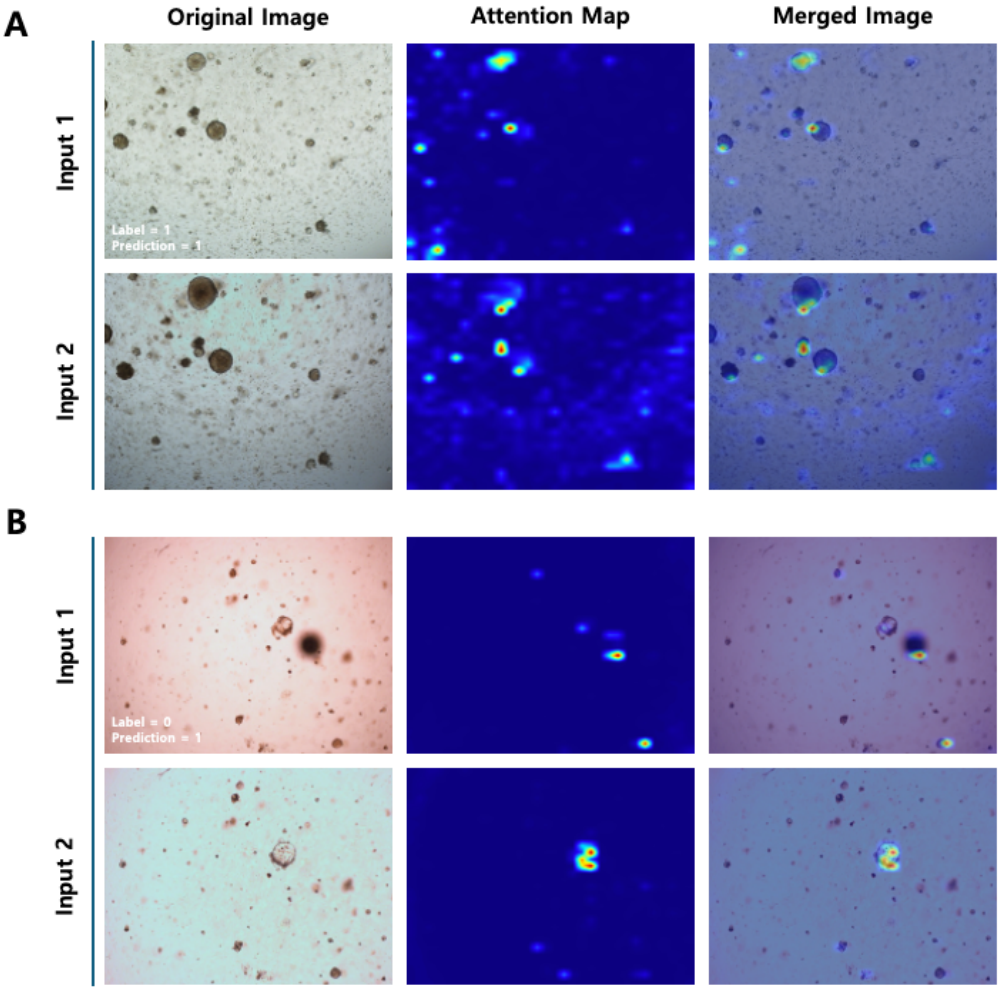
In summary, both models performed well for the HER2 and TNBC subtypes, with the ViT-Dual-MAE-Linearprobe Model consistently showing better performance, particularly for the Luminal A subtype. The Luminal B subtype remains challenging for both models.

**Table 3. Model Prediction Performance Across Breast Cancer Subtypes**

| Subtype | ViT-Dual-MAE Model | Accuracy | Precision | Recall | F1 Score | Number of Image pairs | Number of Patients |
|---|---|---|---|---|---|---|---|
| HER2 | Linearprobe | 90.27 | 93.28 | 92.55 | 92.91 | 370 | 3 |
| | Finetune | 90.27 | 92.61 | 93.33 | 92.97 | | |
| Luminal A | Linearprobe | 89.35 | 97.70 | 89.95 | 93.66 | 432 | 5 |
| | Finetune | 75.46 | 91.72 | 79.10 | 84.94 | | |
| TNBC | Linearprobe | 83.10 | 98.56 | 81.47 | 89.21 | 491 | 3 |
| | Finetune | 85.77 | 90.72 | 92.87 | 91.78 | | |
| Luminal B | Linearprobe | 76.81 | 54.80 | 100.00 | 70.80 | 345 | 3 |
| | Finetune | 69.28 | 47.72 | 96.91 | 63.95 | | |

### 3.9. Analysis of attention maps from the ViT-Dual-MAE-Finetune model

The analysis of attention maps from the ViT-Dual-MAE-Finetune model provided insights into how the model focuses on specific features in organoid microscopy images. When predictions matched actual successful cases, the model often concentrated on specific edges of individual organoids (Figure 7A). This suggests that the model might use these morphological boundaries to assess success. Additionally, some attention was scattered across the areas surrounding the main organoids, indicating that the model may also consider the broader context of each organoid. In cases where the model incorrectly predicted failures as successes, the focus tended to be on the edges or inside areas of the individual organoids, but attention to surrounding structures was generally minimal, which may lead to occasional misclassifications (Figure 7B). For actual failures that the model correctly identified, attention was less commonly placed directly on the organoids themselves and more on their surrounding areas (Figure 7C). This pattern suggests that the model might use features around the main organoids to identify failing conditions. In cases where actual successes were mistakenly classified as failures—often occurs in Passage 0 conditions—the inclusion of various tissues and debris in the gel matrix created dark microscopy images (Figure 7D). These unusual visual characteristics might sometimes cause the model to misinterpret them as indicators of failure.



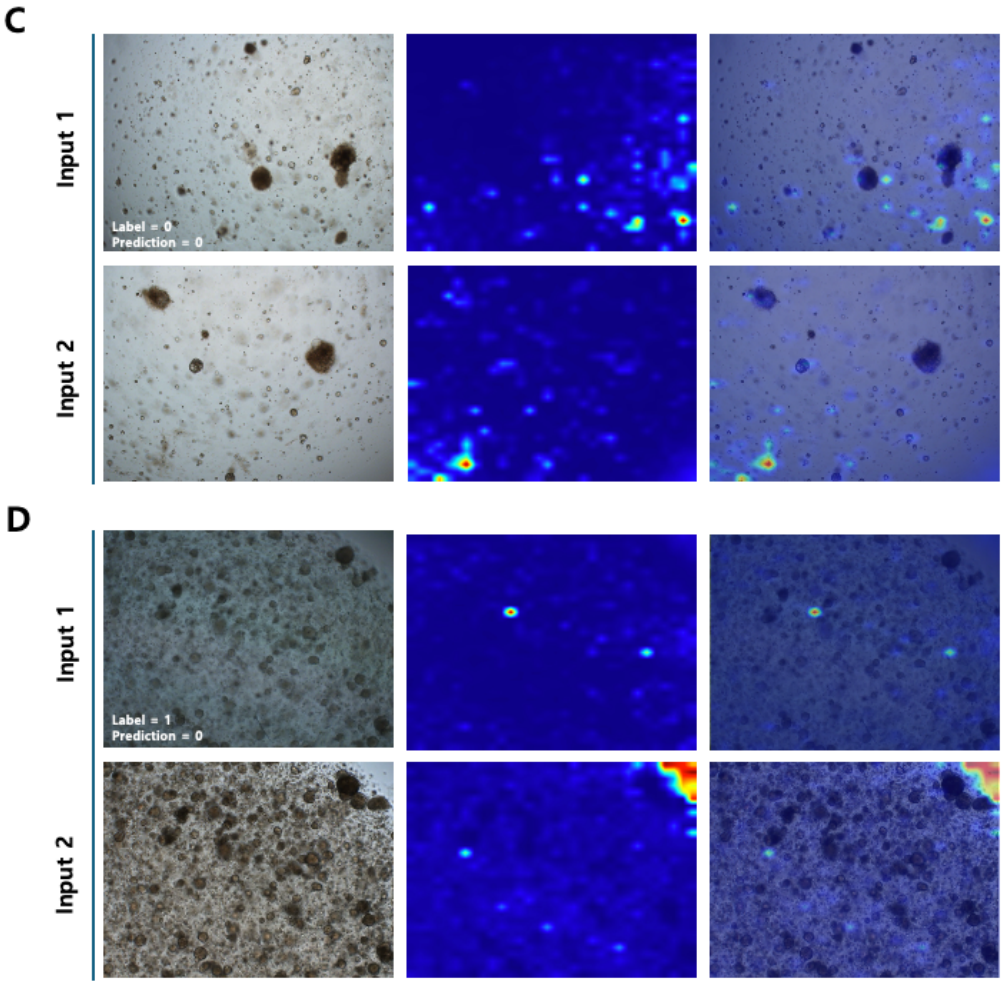



**Figure 7. Analysis of attention map between image pairs.** Original images, attention maps, and combined images with 50% transparency overlay are shown for paired images. Input 1 and 2 are paired images. (A) represents a successful case correctly predicted as successful, while (B) represents a failed case incorrectly predicted as successful. (C) represents a failed case correctly predicted as failed, and (D) represents a successful case incorrectly predicted as failed. In original images, 'Label' indicates the actual result, showing whether the organoid was successful or failed. 'Prediction' refers to the model's predicted outcome, with '1' denoting success and '0' denoting failure.

## 3.10. Time consumption in failed cases

An analysis of 37 unsuccessful organoid cases was conducted to understand the optimal timing for early predictions and how these predictions can be most effectively applied to optimize resource use. The distribution of cases across different passages revealed a higher concentration of failures in the early stages, with Passage 0 having the most failures (11 cases), followed by Passages 1 and 2, each with 7 cases (Figure 8A). Further examination of organoids that reached Passage 4 or beyond revealed instances where initial growth rates appeared normal but eventually exhibited significantly slower growth as the passages progressed. These cases resulted in extended culture periods based on researchers' judgment. Overall, the duration analysis of all failed cases showed an average passage reach of 1.97.

The average time required for each passage varied, with Passage 0 averaging 21.17 days, Passage 1 requiring 36.16 days, and Passage 2 taking up to 37.33 days (Figure 8B). An ANOVA test conducted to evaluate the statistical significance of mean differences among these groups revealed that, excluding samples from Passage 5 and beyond (due to low sample counts of four or fewer), the test for Passage 0 through 4 yielded a significant p-value of 0.0037 (Table 4). Tukey HSD test results showed significant differences in duration between Passage 0 and subsequent passages, indicating that organoid growth rate was faster during Passage 0. The average growth rate during Passage 0 was 17.47 days faster compared to Passage 1 (p-value = 0.0011), 18.48 days faster compared to Passage 2 (p-value = 0.001), and 16.13 days faster compared to Passage 4 (p-value = 0.0206).

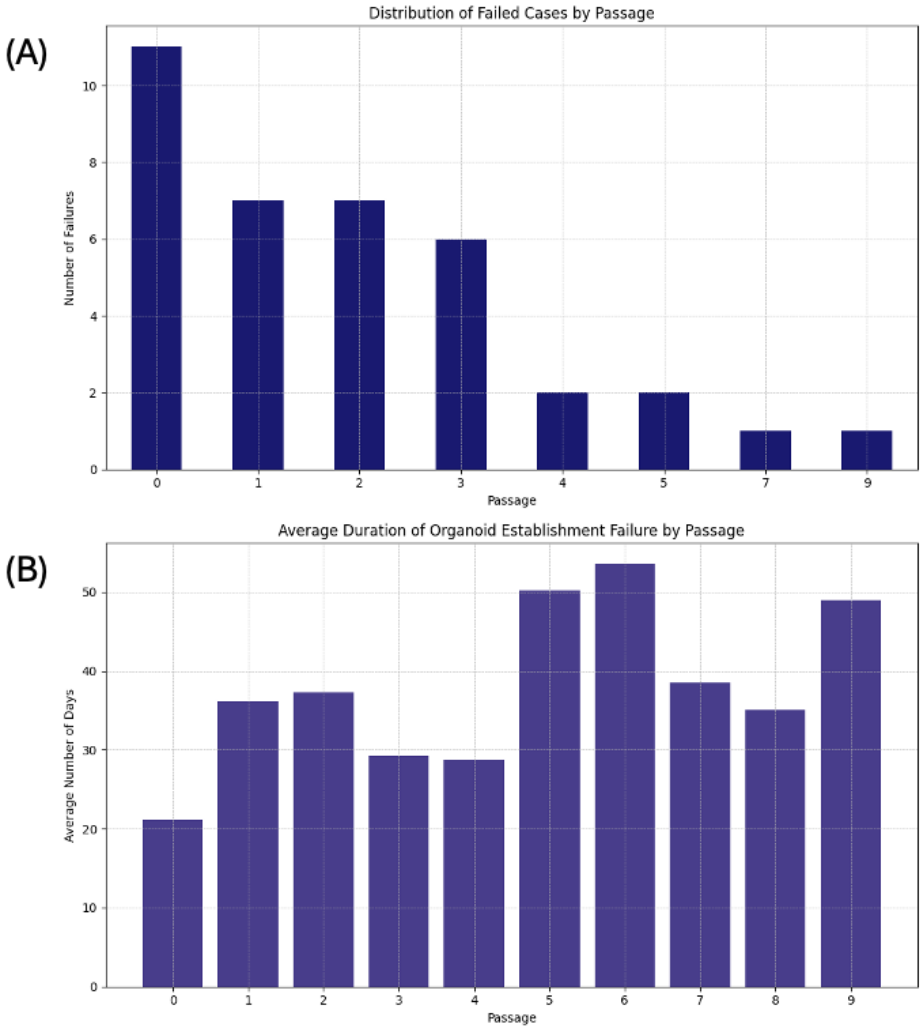



**Figure 8. Duration of Passages and Days in failure cases.** (A) Final passage reached in failure cases are depicted. (B) Average days spent per passage in failure cases are demonstrated.

**Table 4. Difference in duration between passage 0 to 4**

| Passage | | | | |
|---|---|---|---|---|
| Group 1 | Group 2 | Mean Difference | 95% CI (Lower, Upper) | P-value |
| 0 | 1 | 17.4676 | (6.7758, 28.1594) | 0.0011* |
| 0 | 2 | 18.4848 | (6.6862, 30.2835) | 0.001* |
| 0 | 3 | 7.5939 | (-5.513, 20.7009) | 0.6293 |
| 0 | 4 | 16.1333 | (1.8708, 30.3959) | 0.0206* |
| 1 | 2 | 1.0172 | (-12.575, 14.6094) | 0.9 |
| 1 | 3 | -9.8737 | (-23.675, 3.9277) | 0.2392 |
| 1 | 4 | -1.3343 | (-17.298, 14.6294) | 0.9 |
| 2 | 3 | -10.8909 | (-24.692, 2.9103) | 0.1504 |
| 2 | 4 | -2.3515 | (-18.315, 13.6121) | 0.9 |
| 3 | 4 | 8.5394 | (-6.262, 23.3408) | 0.5029 |

*p-value < 0.05 was considered statistically significant.

# 4. Discussion

This study highlights the potential of AI in advancing organoid research by providing a reliable early prediction model for the success of organoid establishment. Our ViT models, particularly the ViT-Dual-MAE-Finetune and ViT-Dual-MAE-Linearprobe models, demonstrated significant promise in accurately predicting the success of organoid cultures. These models not only matched but often surpassed human performance in predicting organoid establishment, which is critical for optimizing resource allocation and enhancing the efficiency of organoid-based research.

The consistent imaging and tracking of organoid establishment provided a robust dataset that allowed for accurate monitoring of morphological changes over time. This comprehensive dataset was essential for training the AI models effectively. The dual-image input approach captured subtle morphological features, improving the models' predictive accuracy[26]. Pretraining the ViT backbone with a MAE further enhanced the models' ability to recognize key morphological indicators, crucial for accurate predictions[25].

The evaluation across different passages showed that the ViT-Dual-MAE-Finetune model excelled at the earliest passage, while the ViT-Dual-MAE-Linearprobe model performed better in subsequent passages. This suggests that different training approaches may be optimal at different stages of organoid development. Notably, the AI models, especially the ViT-Dual-MAE-Linearprobe model, demonstrated higher accuracy, precision, and recall compared to human predictors in later passages, indicating that AI models can effectively complement human analysis, particularly when morphology become more complex[20].

The performance of the models across different breast cancer subtypes varied, reflecting the distinct morphological characteristics of each subtype. The models showed high accuracy and precision for HER2 and TNBC subtypes but faced challenges with the Luminal B subtype, necessitating further refinement to handle diverse morphological features. Luminal B breast cancer is known for its high cell proliferation with elevated Ki-67 levels[27]. The variability in hormone receptor expression and the potential presence of HER2 further contribute to the complexity of this subtype[28]. Consequently, the current organoid establishment protocol, which uses the same media for all subtypes, may result in varied responses for Luminal B, potentially reducing prediction accuracy for this subtype.

The analysis of attention maps provided insights into the features the models focused on during prediction. Successful predictions often concentrated on the edges of individual organoids, suggesting that the edge morphology might be an important indicator of organoid growth and success. Studies on colon organoids have shown that edge features, such as crypt formation and opacity, are crucial indicators of organoid maturity[21]. Additionally, in epithelial organoids, edge morphology significantly contributes to growth dynamics[29]. As the characteristics of colon and epithelial organoids differ from those of breast cancer organoids, it is essential to conduct further research on breast cancer organoids to understand the role of organoid edges.

There are certain limitations to this model. Currently, the application of this model is limited to breast cancer organoids, and its potential for generalization to other types of cancer organoids remains to be tested. Understanding the specific morphological features used by the model for prediction is still limited, which can be a barrier to fully understanding the biological relevance of the model's predictions. In the field of medical artificial intelligence, efforts must be made to uncover the underlying factors in AI decisions using tools like attention maps. Attention maps and other explainability techniques are crucial for interpreting deep learning models and making their decision processes more transparent especially in biological and clinical settings[30,31].

In practical applications, our model demonstrates considerable potential as a decision-support tool in large-scale projects. While the current performance may be insufficient for critical decision-making on individual tissue samples, it offers significant advantages in resource allocation and time management. As highlighted in Result 9, our analysis of time consumption in failed cases indicates that achieving accurate predictions before organoid passages reach the average of approximately two can substantially reduce time investment. This early predictive capability allows researchers to efficiently concentrate efforts on the most promising organoid cultures, minimizing the time and costs associated with maintaining less viable samples.

In conclusion, our study underscores the potential of AI models that predict the success of organoid establishment for organoid based research. While our ViT-Dual-MAE models demonstrated high accuracy and robustness, further refinement is needed to address performance variability across subtypes and extend applicability to other cancer types. Understanding key morphological features and employing explainability tools like attention maps will enhance model transparency and reliability. Ongoing research in this area will advance personalized cancer treatments, highlighting the vital role of AI in precision medicine.

# 5. Conclusion

This study underscores the significant potential of AI-based analysis in the early prediction of organoid establishment success using microscopic images. Utilizing a dual-ViT architecture, the AI model demonstrated high accuracy in detecting subtle morphological features, allowing for reliable predictions of organoid success. The model's performance at early stages highlights its utility in optimizing resource allocation by identifying the most promising organoid cultures. This capability addresses the high failure rates in organoid research, enhancing efficiency and scalability. Integrating advanced AI techniques into organoid research paves the way for significant advancements in precision medicine. The early prediction model supports more efficient drug discovery and the development of personalized treatment strategies, contributing to better patient outcomes. In conclusion, the AI model's ability to accurately predict organoid establishment success early on can transform organoid research, making it more efficient and scalable. This advancement is a crucial step toward enhancing personalized medicine and developing more effective cancer treatments.

## References

1. Turashvili G, Brogi E. Tumor Heterogeneity in Breast Cancer. Front Med (Lausanne) 2017;4:227.
2. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000;406:747-52.
3. Jiang YZ, Ma D, Suo C, Shi J, Xue M, Hu X, et al. Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. Cancer Cell 2019;35:428-40 e5.
4. Morgan MM, Johnson BP, Livingston MK, Schuler LA, Alarid ET, Sung KE, et al. Personalized in vitro cancer models to predict therapeutic response: Challenges and a framework for improvement. Pharmacol Ther 2016;165:79-92.
5. Bleijs M, van de Wetering M, Clevers H, Drost J. Xenograft and organoid model systems in cancer research. EMBO J 2019;38:e101654.
6. Chitrangi S, Vaity P, Jamdar A, Bhatt S. Patient-derived organoids for precision oncology: a platform to facilitate clinical decision making. BMC Cancer 2023;23:689.
7. Tuveson D, Clevers H. Cancer modeling meets human organoid technology. Science 2019;364:952-5.
8. Pauli C, Hopkins BD, Prandi D, Shaw R, Fedrizzi T, Sboner A, et al. Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine. Cancer Discov 2017;7:462-77.
9. Sachs N, de Ligt J, Kopper O, Gogola E, Bounova G, Weeber F, et al. A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell 2018;172:373-86 e10.
10. Veninga V, Voest EE. Tumor organoids: Opportunities and challenges to guide precision medicine. Cancer Cell 2021;39:1190-201.
11. Zhao Y, Huang ML, Zhu YF, Xu LX, Li XX, Yu J. Patient-Derived Organoids Can Predict Chemotherapy Response of Gastric Cancers and Analysis of Its Molecular Characteristics. Gut 2022;71:A47-A8.
12. 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;159:176-87.
13. Mou Y, Huang J, Yang W, Wan Y, Pu Z, Zhang J, et al. Patient-derived primary breast cancer cells and their potential for predicting sensitivity to chemotherapy. Front Oncol 2022;12:1023391.
14. Mazzucchelli S, Piccotti F, Allevi R, Truffi M, Sorrentino L, Russo L, et al. Establishment and Morphological Characterization of Patient-Derived Organoids from Breast Cancer. Biol Proced Online 2019;21:12.
15. Magre L, Verstegen MMA, Buschow S, van der Laan LJW, Peppelenbosch M, Desai J. Emerging organoid-immune co-culture models for cancer research: from oncoimmunology to personalized immunotherapies. J Immunother Cancer 2023;11.
16. Ooft SN, Weeber F, Dijkstra KK, McLean CM, Kaing S, van Werkhoven E, et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. Sci Transl Med 2019;11.
17. Kim M, Mun H, Sung CO, Cho EJ, Jeon HJ, Chun SM, et al. Patient-derived lung cancer organoids as in vitro cancer models for therapeutic screening. Nat Comm

un 2019;10:3991.
18. Beck LE, Lee J, Cote C, Dunagin MC, Lukonin I, Salla N, et al. Systematically quantifying morphological features reveals constraints on organoid phenotypes. Cell Syst 2022;13:547-60 e3.
19. Kassis T, Hernandez-Gordillo V, Langer R, Griffith LG. OrgaQuant: Human Intesti nal Organoid Localization and Quantification Using Deep Convolutional Neural Net works. Sci Rep 2019;9:12479.
20. Matthews JM, Schuster B, Kashaf SS, Liu P, Ben-Yishay R, Ishay-Ronen D, et al. OrganoID: A versatile deep learning platform for tracking and analysis of single-organoid dynamics. PLoS Comput Biol 2022;18:e1010584.
21. 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;22:4118-28.
22. Borten MA, Bajikar SS, Sasaki N, Clevers H, Janes KA. Automated brightfield m orphometry of 3D organoid populations by OrganoSeg. Sci Rep 2018;8:5319.
23. Park T, Kim TK, Han YD, Kim KA, Kim H, Kim HS. Development of a deep l earning based image processing tool for enhanced organoid analysis. Sci Rep 202 3;13:19841.
24. Chefer H, Gur S, Wolf L. Transformer Interpretability Beyond Attention Visualizati on. 2021 Ieee/Cvf Conference on Computer Vision and Pattern Recognition, Cvpr 2021 2021:782-91.
25. He KM, Chen XL, Xie SN, Li YH, Dollár P, Girshick R. Masked Autoencoders Are Scalable Vision Learners. 2022 Ieee/Cvf Conference on Computer Vision and Pattern Recognition (Cvpr 2022) 2022:15979-88.
26. Chicco D. Siamese Neural Networks: An Overview. Methods Mol Biol 2021;2190: 73-94.
27. Lopez-Gonzalez L, Sanchez Cendra A, Sanchez Cendra C, Roberts Cervantes ED, Espinosa JC, Pekarek T, et al. Exploring Biomarkers in Breast Cancer: Hallmarks of Diagnosis, Treatment, and Follow-Up in Clinical Practice. Medicina (Kaunas) 2 024;60.
28. Tran B, Bedard PL. Luminal-B breast cancer and novel therapeutic targets. Breast Cancer Res 2011;13:221.
29. Hof L, Moreth T, Koch M, Liebisch T, Kurtz M, Tarnick J, et al. Long-term live imaging and multiscale analysis identify heterogeneity and core principles of epithe lial organoid morphogenesis. BMC Biol 2021;19:37.
30. Srinivasu PN, Sandhya N, Jhaveri RH, Raut R. From Blackbox to Explainable AI in Healthcare: Existing Tools and Case Studies. Mobile Information Systems 2022; 2022.
31. Yang G, Ye Q, Xia J. Unbox the black-box for the medical explainable AI via m ulti-modal and multi-centre data fusion: A mini-review, two showcases and beyond. Inf Fusion 2022;77:29-52.

# APPENDICES

| | |
|---|---|
| PDO | : patient derived organoid |
| MAE | : masked autoencoder |
| ViT | : vision transformer |
| AUC | : area under curve |
| AI | : artificial intelligence |
| GUI | : graphical user interface |
| MLP | : multi-layer perceptron |
| ROC | : receiver operating characteristic |
| HER2 | : human epidermal growth factor receptor 2 |
| TNBC | : triple negative breast cancer |

Abstract in Korean

# 딥러닝 기반 현미경 이미지 분석을 통한 오가노이드 구축 성공의 조기 예측

본 연구는 환자 유래 오가노이드 구축 조기 단계에서 구축 성공 여부를 예측하는 인공지능 모델을 개발하여 구축의 효율성을 향상시키는 것을 목표로 한다. 이러한 예측은 성공할 것으로 예상되는 오가노이드 구축에 전략적으로 자원을 재분배하여 집중된 연구를 가능하게 하며 정밀 의학의 발전에 기여한다. 데이터로 사용된 오가노이드 현미경 사진은 유방암 오가노이드 구축과정에서 촬영되었다. 같은 오가노이드의 같은 계대에서 서로 일정 시간 간격을 갖는 현미경 사진이 쌍을 이루어 두개의 비전 트랜스포머 구조의 딥 러닝 모델에 제공된다. 이 접근 방식은 모델이 시간에 따른 오가노이드의 미세한 형태학적 변화를 감지할 수 있도록 한다. 이미지 쌍을 제공하는 방식과 Masked Autoencoder를 사용한 사전학습을 결합한 비전 트랜스포머 모델은 높은 예측 정확도를 보였으며, 여러 지표에서 인간의 예측과 유사하거나 우수한 성능을 보였다. ViT-Dual-MAE-Finetune 및 Linearprobe 모델이 특히 우수한 성능을 보였으며, 0.88의 AUC를 달성하였다. ViT-Dual-MAE-Linearprobe 모델은 오가노이드 계대가 진행됨에 따라 성능이 향상되었으며, 계대 2에서 AUC 0.98을 기록했다. 또한 유방암 아형에 따른 성능 분석에서 HER2와 TNBC에서 우수한 성능을 보였다. 어텐션 맵을 분석한 결과, 실제로 성공인 케이스를 잘 예측한 경우 종종 각 오가노이드의 가장자리에 어탠션이 집중되는 경향을 확인하였으며, 이는 이러한 특징이 오가노이드의 성장과 구축 성공여부를 파악하는데 중요한 지표일 수 있음을 시사한다. 정밀 의학 및 약물개발 분야에서 인공지능 기반 예측 모델을 활용한 오가노이드 구축 여부의 조기 결정은 제한된 자원의 최적화된 사용을 가능하도록 하는 중요한 접근법으로 사용될 수 있다.

---

**핵심되는 말** : 환자 유래 오가노이드, 인공지능, 구축 조기 예측