

Biological Validation of Attribute Classification of Cell Nuclei in Three-dimensional Myocardial Tissue Images of Mice

Takamitsu Araki,
Graduate School of Science and
Technology,
Kumamoto University,
Kumamoto, Japan
araki@st.cs.kumamoto-u.ac.jp

Yuichiro Arima,
Developmental Cardiology Laboratory,
International Research Center for
Medical Sciences,
Kumamoto University,
Kumamoto, Japan
arimay@kumamoto-u.ac.jp

Masahiro Migita
Research and Education Institute for
Semiconductors and Informatics,
Kumamoto University,
Kumamoto, Japan
migita@cc.kumamoto-u.ac.jp

Masashi Toda,
Research and Education Institute for
Semiconductors and Informatics,
Kumamoto University,
Kumamoto, Japan
toda@cc.kumamoto-u.ac.jp

Abstract—Heart disease is the second leading cause of death in Japan, accounting for 15% of all deaths. Since there are limited treatment options for severe heart disease, there is a need to develop new treatment methods. To this end, it is necessary to analyze the heart at the cellular level and elucidate the pathogenesis of the disease. We have developed a system to calculate cellular composition ratios from 3D myocardial images of mice. However, the number of samples analyzed is still small, so verification of generality is an issue. In this study, we analyzed new samples and improved the classification method. In addition to the cell composition ratios, we also focused on the degree of cell division by introducing a new analytical index. The results obtained were biologically consistent in many respects, confirming the validity of the current method.

Keywords—heart disease, knock-out mice, image analysis, cell classification, support vector machine

I. INTRODUCTION

Cardiac disease accounts for about 15% of all deaths in Japan, and is the second leading cause of death after malignant neoplasms [1]. Heart failure accounts for about 41% of all deaths from heart disease [2]. Heart failure is treated with drugs, catheterization, and bypass surgery [3], but the only established treatment for severe cases of heart failure is heart transplantation [4]. However, Japan has a strict organ transplant guideline system, so the number of organ transplants in Japan is far fewer than in other countries [5].

One of the reasons for the limited treatment options for severe heart failure is that once myocardial cells are damaged, they do not regenerate spontaneously. However, there is an exception to this, and it is known that cardiomyocytes can divide during the early postnatal period. Therefore, the analysis of the hearts of newborn animals is the starting point for the search for new treatment methods.

Recent advances in imaging and computer technology have made it possible to analyze the heart at the cellular level. A confocal microscope are one of the instruments used to image cells. In addition to the ability to obtain high-contrast, high-resolution three-dimensional images, confocal microscopes have the advantage of being easier to handle than other microscopes. Basically, four fluorescence sites can be selected, and each can be output separately. Although there is

a large amount of information obtained from 3D images, there are problems such as the need for specialized knowledge and a large human workload required for analysis. Therefore, we developed a system to count cell nuclei by type using machine learning on 3D myocardial images of mice[6]. However, the number of data actually analyzed by this method was small and lacked generality. In this study, we analyze additional samples using the developed cell classification method and verify its validity from a biological perspective. For this purpose, we added additional analytical indices and improved the algorithm.

II. PREVIOUS WORK

We have developed a method for cell nucleus classification using 3D myocardial images of mice taken with a confocal microscope. Specifically, we calculated the features that reflect the biological characteristics and morphological information of each cell nucleus, and classified them into three classes: cardiac myocyte nuclei, vascular endothelial cell nuclei, and other cell nuclei. Figure 1 shows the flow of the classification method. As a comparison, image classification of cell nuclei using deep learning is performed, in which each cell nucleus is cut out from the original image using a bounding box and classified using a 3D CNN. We compared the classification results of this method with those of a support vector machine using the proposed features. The results showed that the proposed method outperformed all the methods in Precision, Recall, and F1-score.

A. Classification Indicators

There are various types of cells that make up the heart, including cardiomyocytes, vascular endothelial cells, smooth muscle cells, and so on. Biological characteristics are the unique features of these cell types. In this study, two biological characteristics are quantified and used for classification: the first is that cardiomyocytes have a smaller proportion of nuclei in the total cell volume than non-cardiomyocytes. To quantify this, we define the filling ratio P_F , which is expressed by the following equation:

$$P_F = \frac{V_n}{V_c} \times 100, \quad (1)$$

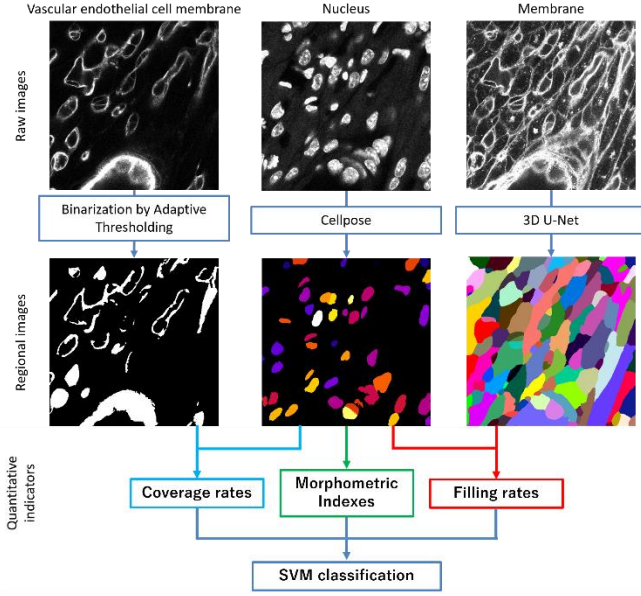


Fig. 1. Previous research approach, classifying cell nuclei on the basis of quantitative indicators: coverage rates, filling rates, various types of morphometric indexes. They are computed from three regional images.

where V_n is the volume of the cell nucleus, and V_c is the volume of the whole cell. Since the cell nucleus and cell region images are separate images, they are mapped to each other using the center-of-gravity coordinates.

The second biological characteristic is that the vascular endothelial cell nucleus is tightly covered by the vascular endothelial cell membrane. To quantify this, the coverage rate P_F is defined and expressed by the following equation:

$$P_c = \frac{n}{m} \times 100, \quad (2)$$

where m is the number of voxels on the surface of the cell nucleus, and n is the number of those voxels adjacent to the vascular endothelial cell membrane.

The morphological information includes volume and surface area, mean curvature, Gaussian curvature, “shape index,” and “curvedness.” [7] The mean curvature H and Gaussian curvature G are expressed by the minimum curvature κ_1 and maximum curvature κ_2 as follows:

$$H = \frac{\kappa_1 + \kappa_2}{2}, \quad (3)$$

$$G = \kappa_1 \times \kappa_2 \quad (4)$$

The shape index and curvedness are indices that quantify the degree of surface irregularity and the amount of curvature, respectively. The shape index SI and curvature CV are expressed by the following equations:

$$SI = \frac{2}{\pi} \arctan \frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2} \quad (5)$$

$$CV = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}} \quad (6)$$

B. Experiment

We compared the results of classification using a support vector machine with the introduced features and image classification using a 3D CNN. Because the number of cell nuclei of each species is not uniform, stratified 5-segment cross-validation was used to calculate an accuracy index.

Table I shows the specifications of the images taken. The images used are microscopic images of the heart of a 7-day-old knock-out (KO) mouse, a mouse whose metabolism is inhibited by preventing the production of HMG-CoA Synthase 2, an enzyme important for the synthesis of ketone bodies [8]. On the other hand, mice that have not been genetically engineered are called wild-type (WT) mice.

The microscopic images used in the experiment consisted of four channels in which the cell nucleus, cardiomyocyte nucleus, cell membrane, and vascular endothelial cell membrane each fluoresced. Therefore, the correct data were generated visually as follows: those fluorescing in the cell nucleus and cardiomyocyte nucleus channels were classified as cardiomyocyte nuclei, those covered by the vascular endothelial cell membrane were classified as vascular endothelial cell nuclei, and all other cell nuclei were classified as other cell nuclei.

For cell nucleus classification using 3D CNN, each cell nucleus was trimmed from the original image based on the mask image of all cell nuclei extracted by Cellpose [9]. There were 834 myocardial cell nuclei, 944 vascular endothelial cell nuclei, and 481 other cell nuclei, and data expansion was performed to add randomly rotated nuclei by 5° , 10° , and 20° counterclockwise or clockwise.

C. Result

Table II shows the F1 scores for each cell type for the SVM and 3D CNN-based classification methods. For simplicity, Precision and Recall values are omitted. The last row shows the average of the three classes. Since SVM is higher for all indices, the proposed method is effective for cell nucleus classification. However, even with the proposed method, the scores of other cells are lower than those of the other two cell types.

III. CHANGES FROM THE PREVIOUS WORK

In this chapter, we describe the changes we made from previous studies in analyzing the new data. There are three changes: the analysis sample was changed, the interrupted cell nuclei were removed, and the dividing cell nuclei were extracted.

TABLE I. SPECIFICATIONS OF IMAGES USED

Format	Tiff
Size of one voxel	0.1803752×0.1803752×0.2985004 (micron)
Resolution	1024×1024×169 (voxel)
Color type	Gray scale
Object	Neonatal mouse (seven days old)
Photographed area	Left atrium
Fluorescent point	All nucleus, myocardial cell nucleus, vascular endothelial cell nucleus, vascular endothelial cell membrane
Equipment name	Leica SP8 confocal microscope

TABLE II. F1 SCORES FOR EACH CELL TYPE FOR BOTH METHODS

Cell type	Method	
	<i>SVM</i>	<i>3D CNN</i>
Cardiomyocytes	0.960	0.926
Vascular Endothelial cells	0.901	0.778
Other cells	0.701	0.510
Mean	0.854	0.738

A. Samples to be Analyzed

The specifications of the image data used in this study are the same as those shown in Table I in Section II.B, except for the fluorescence locations. The fluorescent areas consist of four channels: cell nucleus, cell membrane, dividing cell nucleus, and vascular endothelial cell membrane. Table III shows a breakdown of the mice photographed. There are six images of postnatal day 3 mice, three each of KO and WT mice. The number of images of mice on the seventh day of life is 17, including 9 images of KO mice and 8 images of WT mice. Mice of different postnatal ages were prepared to see if the cell nuclear classification results reflect biological differences between the two groups. For example, since it is biologically known that cardiomyocytes account for a higher percentage of total cells in 3-day-old mice than in 7-day-old mice, the classification method can be validated from this perspective.

B. Removal of Broken Cell Nuclei

One of the challenges of the developed classification method was that many of the extracted cell nuclei were very small. Figure 2 shows the volume distribution of cell nuclei in a single image of the 3-day-old mouse, and the number of cell nuclei is counted every $20 \mu\text{m}^3$. This figure shows that there are many cell nuclei with a volume of $0 \mu\text{m}^3$ to $20 \mu\text{m}^3$. However, such small nuclei do not generally exist, so the cause of the problem must be investigated and remedied.

The investigation revealed that these nuclei are nuclei that are broken off at the edges of the image. Such nuclei must be removed because the feature calculations are affected when cell nuclei are cut off. Figure 3 shows the volume distribution after removing the broken nuclei. It can be seen that most of the cell nuclei between $0 \mu\text{m}^3$ and $40 \mu\text{m}^3$, which were the problem, have been removed.

C. Extraction of Cell Nuclei during Division

In this study, we also extracted and classified cell nuclei in division. The purpose was to quantify how active division is in each cell type. The facts that can be verified by this study are that cardiomyocytes divide actively in postnatal day 3 mice, but the rate of division declines rapidly by day 7, while vascular endothelial cells divide constantly. There are several phases of cell division, including G1, S, G2, and M phase. The nuclei of cells in the G2 and M phase are photographed under the microscope, but the nuclei of cells in the M phase are the subject of this study. Figure 4 shows the channel of dividing cell nuclei, and the apparent difference between G2 and M-phase nuclei is that the G2-phase nuclei fluoresce sparsely, while the M-phase nuclei fluoresce strongly throughout the nucleus.

TABLE III. SAMPLE BREAKDOWN

	KO mice	WT mice	Total
three days old	3	3	6
seven days old	9	8	17

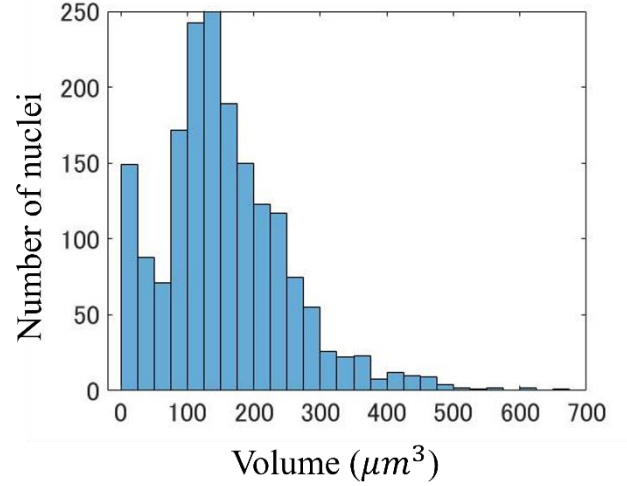


Fig. 2. Volume distribution before removing interrupted nuclei. There are many very small cell nuclei, which is a biologically impossible situation.

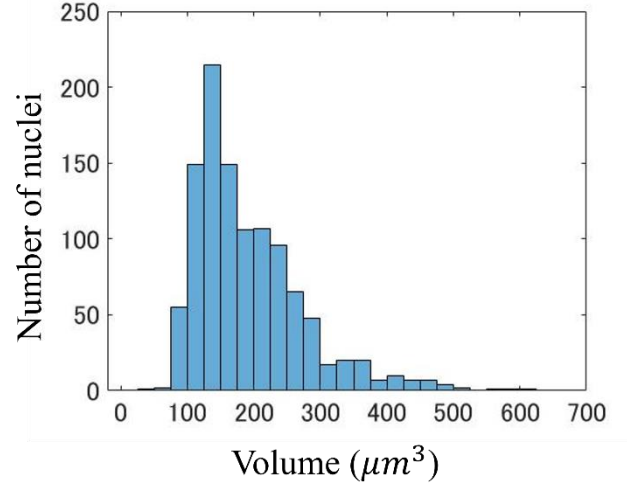


Fig. 3. Volume distribution after removing interrupted nuclei. Mainly very small cell nuclei were removed, resulting in a highly valid volume distribution.

To extract the M phase, the original image was first binarized to remove noise. Specifically, the connected components with a volume of less than 3000 voxels were removed. However, at this point, the G2 phase still remains, and the cell nucleus image has not been label-matched with the segmentation result. Therefore, the G2 and M phase region images were superimposed on the cell nucleus region image to identify the M phase cell nucleus region based on the degree of overlap. The overlapping ratio α is expressed by the following equation for each cell nucleus:

$$\alpha = \frac{m}{n}, \quad (7)$$

where m is the number of voxels constituting each nucleus and n is the number of those voxels in the G2 and M phase regions. Figure 5 shows the number of cell nuclei by overlapping ratio for sample 2 of postnatal day 3 KO mice. Although the maximum value on the vertical axis is restricted to 30, the leftmost interval contains more than 1000 nuclei. These correspond to nuclei that are not dividing, while highly

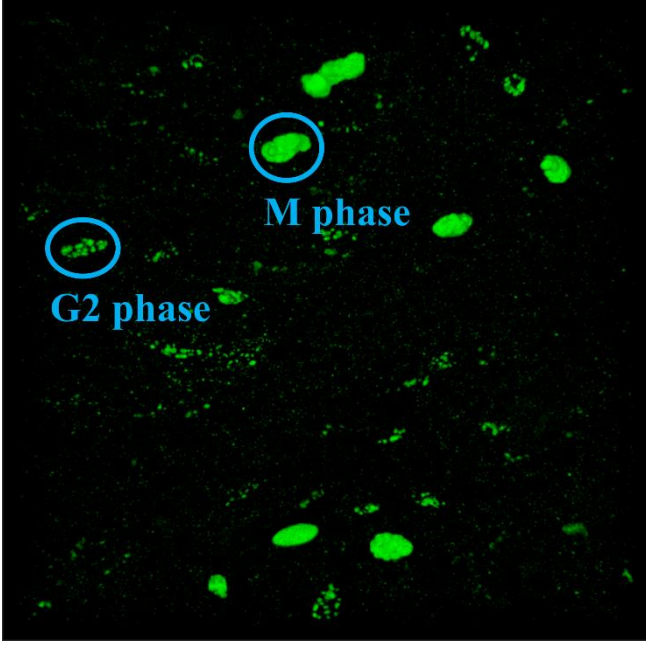


Fig. 4. Microscopic image of cell nuclei in division. The M stage is characterized by strong fluorescence throughout, while the G2 stage is characterized by sparse fluorescence. There is also a lot of noises.

overlapping nuclei correspond to M-phase nuclei and moderately overlapping nuclei correspond to G2-phase nuclei.

IV. EXPERIMENT

In the experiment, we did two things. First, we removed the broken cell nuclei for the two images used in Section II.B and compared the accuracy with that shown in Table II. Next, after removing the cut off cell nuclei, cell nucleus classification was performed on the 23 images described in Section III.A. In this process, cell nuclei in the M phase were also extracted. SVM was used for classification, and the features used were the same as those described in Section II.A. The two images used in Section II.B were used as the training dataset.

To quantify the cellular composition ratio of mice, we define the myocardial rate P_{CM} as follows:

$$P_{CM} = \frac{n_{CM}}{n_{ALL}} \times 100, \quad (8)$$

where n_{CM} is the number of myocardial cell nuclei in each sample image and n_{ALL} is the number of all cell nuclei in each sample image. The same calculation method is applied to vascular endothelial cell nuclei and other cell nuclei to obtain the vascular endothelial rate and other cell rate, respectively.

To quantify the activity of cell division, we define the dividing rate D for each cell type in each sample image as follows:

$$D = \frac{n_D}{n_{all}} \times 100, \quad (9)$$

where n_D is the number of cell nuclei in M phase for each cell type in each sample image. n_{all} is the number of cell nuclei for each cell type in each sample image. This calculation method is applied to myocardial cell nuclei, vascular endothelial cell nuclei, and other cell nuclei to obtain

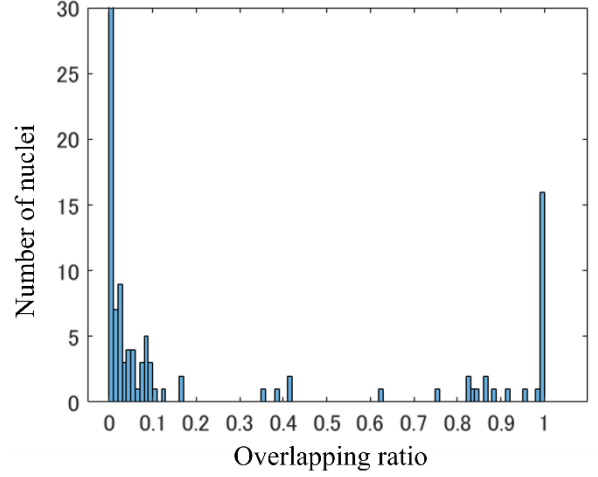


Fig. 5. Distribution of overlapping ratios of cell nuclei. The upper limit of the vertical axis is 30. The degree of overlap indicates the degree of cell division activity, so a nucleus close to 0 indicates a nucleus that is not dividing, while a nucleus close to 1 indicates a nucleus in the M phase.

myocardial dividing rate, vascular endothelial dividing rate, and other dividing rate, respectively.

V. RESULT

In this section, we first show the change in accuracy due to the removal of the broken cell nuclei. Next, we present the results of the cell nucleus classification. The case of all cell nuclei and the case of only M-phase cell nuclei are considered, and evaluated using the indices described in Section IV. For the sake of simplicity, the values for each sample are omitted and only the mean values are given. Mean values were calculated in four ways: on the third or seventh postnatal day, WT or KO mice. In the following tables, “CM,” “VE,” and “OT” are used as an abbreviation for cardiomyocyte nuclei, vascular endothelial cell nuclei and other cell nuclei respectively.

A. Accuracy Change due to Removal of Broken Cell Nuclei

Table IV compares the F1 scores with and without removal of the broken cell nuclei. The sample images and classification method used are exactly the same as in Section II. All cell types show improved scores, with particularly high increases for vascular endothelial cells and other cells.

B. Classification Results of All Cell Nuclei

Table V shows the number of cell nuclei per type. Table VI shows the results of the calculation of the composition ratios of each cell type using the values in Table V. Three days old WT mice had a 5% higher myocardial rate and 1% lower vascular endothelial rate at 3 days compared to 7 days old KO mice. Three days old KO mice had a 20% higher myocardial rate and 17% lower vascular endothelial rate. As for the KO mice, myocardial rate was 20% higher and vascular endothelial rate was 17% lower on day 3. This is consistent

TABLE IV. F1 SCORES BEFORE AND AFTER REMOVAL OF THE DISCONNECTION

Cell type	Method	
	With Removal	Without Removal
Cardiomyocytes	0.963	0.960
Vascular Endothelial cells	0.921	0.901
Other cells	0.741	0.701
Mean	0.901	0.854

TABLE V. NUMBER OF CELL NUCLEI BY TYPE

	CM	VE	OT	Total
7-day-old KO	396	430	110	936
7-day-old WT	399	356	126	880
3-day-old KO	593	275	92	959
3-day-old WT	395	327	61	783

TABLE VI. PROPORTION OF CELL NUCLEI BY TYPE

	CM	VE	OT
7-day-old KO	42	46	12
7-day-old WT	45	41	14
3-day-old KO	62	29	10
3-day-old WT	50	42	8

with the fact that myocyte division declines as postnatal days pass, while vascular endothelial cell division remains constant [10]. In addition, comparing KO and WT mice on day 3, the myocardial rate was 12% higher and the vascular endothelial rate was 13% lower in the KO mice. This also seems to correctly reflect the nature of the KO mice.

C. Classification Results of Cell Nuclei in M Phase

Table VII shows the mean values for the classification of cell nuclei in M phase. Table VIII shows the results of the dividing rates calculation using these values in Table VII. First, the dividing rates on postnatal day 7 was compared between KO and WT mice. For all cell types, the dividing rate of KO mice is higher than that of WT mice. This may be related to the characteristics of KO mice described in Section II.B. That is, KO mice have a slow progression of cell division due to metabolic inhibition, and the number of cells that are dividing at a given moment is higher. The dividing rate of WT mice on day 7 was less than 1% for myocardium and slightly above 1% for vascular endothelium. This is very close to the values calculated by other biological methods.

Myocardial dividing rates on day 3 is less than 1% in both cases, whereas the correct value for WT mice is between 5% and 10%. Judging from the original images, this may be due to staining at the time of imaging. Therefore, it is necessary to reanalyze the images taken again.

VI. CONCLUSION

In this study, we validated the cell nucleus classification method that we have been developing. Specifically, we compared 3-day-old and 7-day-old mice to determine whether the classification method reflected biological differences between the two groups. We also extracted dividing cell nuclei and verified the classification from the viewpoint of dividing rate. In addition, we removed the broken nuclei to improve the accuracy of classification. As a result, the validity of the classification method was confirmed from various viewpoints. We will continue to increase the number of samples to obtain highly generalized analysis results.

One disadvantage of this method is the large number of microscope image channels used. Of the four channels that can be used, three channels are required for the cell nucleus, cell membrane, and vascular endothelial cell membrane, which places significant constraints on the imaging process.

TABLE VII. NUMBER OF M-PHASE CELL NUCLEI BY TYPE

	CM	VE	OT	Total
7-day-old KO	4	7	2	13
7-day-old WT	2	4	1	7
3-day-old KO	2	6	2	10
3-day-old WT	0	3	1	5

TABLE VIII. DIVIDING RATE OF M-PHASE CELL NUCLEI BY TYPE

	CM	VE	OT
7-day-old KO	1.0	1.6	1.9
7-day-old WT	0.51	1.2	0.77
3-day-old KO	0.38	2.3	2.2
3-day-old WT	0.067	0.98	1.7

A different approach is the three-dimensional classification of cell nuclei based on deep learning. By using a neural network for feature extraction, the number of occupied channels may be reduced.

REFERENCES

- [1] "Ministry of Health, Labour and Welfare, Summary of the 2020 Vital Statistics Monthly Report Annual Total (Approximate)," [Online]. Available: <https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai20/d1/gaikyouR2.pdf>.
- [2] National Cardiovascular Center, "Epidemiology of Heart Failure and Ischemic Heart Disease," [Online]. Available: https://www.ncvc.go.jp/coronary2/column/20211209_05.html
- [3] New Heart Watanabe International Hospital, "Heart Failure," [Online]. Available: https://newheart.jp/glossary/detail/cardiovascular-surgery_008.php
- [4] Ministry of Education, Culture, Sports, Science and Technology, "Strategic Research on Heart Failure: Pathophysiology and Gene/Cell Therapy of Heart Failure Using Developmental Engineering," [Online]. Available: https://www.mext.go.jp/a_menu/shinkou/hojyo/1300506.htm
- [5] Japan Organ Transplant Network, "Are there any differences in transplant care provided overseas and in Japan?" [Online]. Available: <https://www.jotnw.or.jp/faq/detail.php?id=85>
- [6] S. Kaneko, Y. Arima, M. Migita, M. Toda, "Cell Classification in Myocardial Tissue Images Using Three-dimensional Morphological Information and Physiological Features," Dynamic Image processing for real Application Workshop 2024, unpublished.
- [7] Jan J Koenderink, Andrea J van Doorn, "Surface shape and curvature scales," Image and Vision Computing, vol. 10. Issue 8, 1992, pp.557-564, DOI:[https://doi.org/10.1016/0262-8856\(92\)90076-F](https://doi.org/10.1016/0262-8856(92)90076-F).
- [8] IRCMS - Kumamoto University, "Yuichiro Arima," [Online]. Available: https://ircms.kumamoto-u.ac.jp/research/yuichiro_arima/
- [9] C. Stringer, T. Wang, M. Michaelos, M. Pachitariu, "Cellpose: a generalist algorithm for cellular segmentation," 2020, bioRxiv:10.1101. DOI:10.1101/2020.02.02.931238
- [10] O. Bergmann, S. Zdunek, A. Felker, et al, "Dynamics of Cell Generation and Turnover in the Human Heart," Cell, Volume 161, Issue 7, pp. 1566–1575, June, 2015, DOI:10.1016/j.cell.2015.05.026