

# Revolutionizing Echocardiography: A Comparative Study of Advanced AI Models for Precise Left Ventricular Segmentation

Dong Ok Kim<sup>a</sup>, Minsu Chae<sup>a</sup>, HwaMin Lee<sup>a,\*</sup>

<sup>a</sup> Department of Biomedical Informatics, Korea University College of Medicine, Seoul, 02708, Republic of Korea

Corresponding author: \*hwamin@korea.ac.kr

**Abstract**—Cardiovascular diseases, a leading cause of global mortality, underscore the urgency for refined diagnostic techniques. Among these, cardiomyopathies characterized by abnormal heart wall thickening present a formidable challenge, exacerbated by aging populations and the side effects of chemotherapy. Traditional echocardiogram analysis, demanding considerable time and expertise, now faces overwhelming pressure due to escalating demands for cardiac care. This study addresses these challenges by harnessing the potential of Convolutional Neural Networks, specifically YOLOv8, U-Net, and Attention U-Net, leveraging the EchoNet-Dynamic dataset from Stanford University Hospital to segment echocardiographic images. Our investigation aimed to optimize and compare these models for segmenting the left ventricle in echocardiography images, a crucial step for quantifying key cardiac parameters. We demonstrate the superiority of U-Net and Attention U-Net over YOLOv8, with Attention U-Net achieving the highest Dice Coefficient Score due to its focus on relevant features via attention mechanisms. This finding highlights the importance of model specificity in medical image segmentation and points to attention mechanisms. The integration of AI in echocardiography represents a pivotal shift toward precision medicine, improving diagnostic accuracy and operational efficiency. Our results advocate for the continued development and application of AI-driven models, underscoring their potential to transform cardiovascular diagnostics through enhanced precision and multimodal data integration. This study validates the effectiveness of state-of-the-art AI models in cardiac function assessment and paves the way for their implementation in clinical settings, thereby contributing significantly to the advancement of cardiac healthcare delivery.

**Keywords**—Cardiomyopathy; echocardiograph; deep learning; YOLO; U-Net; attention U-Net.

Manuscript received 11 Oct. 2023; revised 17 Dec. 2023; accepted 22 Jan. 2024. Date of publication 30 Jun. 2024.  
IJASEIT is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.



## I. INTRODUCTION

Cardiovascular diseases have risen to be one of the major causes of death globally, posing an intricate health challenge [1], [2]. Cardiomyopathies of all these diseases, characterized by the abnormal thickening of the heart wall, have come under the prominent medical scanner for raising concerns of high order [3], [4]. This is often further compounded by side effects from chemotherapeutic agents. Without question, this reflects a strong need for accurate diagnostic approaches. Furthermore, demographic changes in populations, such as aging societies, are likely to result in even more direct and exacerbating pressure on the health system about these multifactorial groups of diseases [5], [6], [7].

To elaborate, the traditional methods of echocardiogram analysis require massive involvement concerning time and expertise, where some trainees should specifically work with maximum care and accuracy to provide a reliable result [8],

[9]. The increasing number of patients looking forward to cardiac care and the growing demand for these professionals and healthcare infrastructure are immense pressures, to put it lightly [10], [11]. Above all, therefore, there is a great need to optimize echocardiogram analysis for better efficiency and to prevent this increased likelihood of diagnostic errors.

Recent advancements in artificial intelligence, particularly Convolutional Neural Networks (CNNs), refer to a very dynamic revolutionizing of the paradigms in medical imaging and diagnostics [9], [10], [11], [12]. CNNs place only remarkably fewer operational strains on the medical staff due to the inherent capability of fine recognition of image patterns, which contributes to improved diagnostic accuracy and far surpassing it [13]. In the study, the potential benefits of CNNs are availed through the application of models like YOLO, U-Net, and Attention U-Net, which, in this case, have been modified specifically for use in the segmentation of medical images using the comprehensive EchoNet-

Dynamic echocardiographic dataset from Stanford University Hospital [13].

This study has been designed to focus on the core objectives that ensure a level ground for the cutting-edge computational models that align with the nuanced world of clinical diagnostics. It focuses on how to further develop and refine the deep learning architecture—YOLOv8, U-Net, Attention U-Net—for precise segmentation of the left ventricle in echocardiography image, whereby each model is developed so that it can address the challenges that are on their own intricate in cardiac imaging.

This is an environment that, in the future, deep learning models would improve the model, citing integration with multimodal sources of data [14], [15]. This is meant to enrich the predictive models with more types of data, such as comprehensive patient history, electrocardiogram (ECG) readings, and magnetic resonance imaging (MRI) data, in a forward approach beyond the limitation of traditional echocardiographic images. Such integration of the dissimilar streams of data is said to enhance diagnostic accuracy to a much higher level, giving a clinician an overall picture of conditions that affect the heart [16], [17].

Tapping into this potential may be done using the diagnostics model's part of the suite, such as YOLOv8, U-Net, Attention U-Net, and many others. Automating the segmentation of the echocardiographic images' left ventricular region would substantially ease the operating, practical application of these techniques in the evaluation of cardiac performance. Efficient segmentation is, therefore, the core point to allow quantification of the respective key cardiac parameters: chambers volumes and ejection fractions, from where metrics essential in diagnoses and management of diseases of the heart will be obtained [18].

Thus, this is not a work focused on high-resolution imaging for further cardiac diagnostics; rather, this is work on the transformational potential that automated segmentation techniques may provide. Such methodologies are designed for the ease of the clinician for better diagnosis and assessment of cardiac health [19], [20]. This work contributed highly to the accuracy and effectiveness of the cardiac function test by using state-of-the-art artificial intelligence models that accurately segment the left ventricle. Such advances promise to bolster clinical workflows without increasing the workload of healthcare providers. This underscores one of the main critical contributions of automatic segmentation to the cardiac function assessment; that the development of state-of-the-art AI models is not just possible but pivotal for even better diagnostic accuracy and operational efficacy in the cardiovascular sphere [21].

Such technologies practically apply in the sense that they help simplify and improve the diagnostic procedures that yield better results for the patient [22]. This work substantially exceeds the detailed development and optimization of the state-of-the-art deep learning models for echocardiography; rather, it pioneers efforts in multimodal data integration to foster continued advancement of clinical workflow. The collaborative research efforts are anticipated to significantly advance the integration of artificial intelligence and cardiology significantly, thereby enabling more accurate, efficient, and comprehensive diagnostic approaches for cardiovascular diseases. The rationale of this

study is to pave the way for smarter AI-enabled solutions in the delivery of cardiac healthcare, hence contributing significantly to this field.

## II. MATERIAL AND METHOD

### A. Dataset

The EchoNet-Dynamic is a publicly accessible collection comprising 10,030 echocardiographic videos [13]. These videos, showcasing the apical-4-chamber view, were captured from patients during routine clinical examinations at Stanford University Hospital between 2016 and 2018. To standardize and maintain confidentiality, each video was formatted to a resolution of 112x112 pixels and edited to eliminate textual data and extraneous details outside the primary scanning region [13]. Expert cardiologists annotated each video, providing key measurements such as left ventricular ejection fraction (LVEF), volumes (end-systolic and end-diastolic), and wall motion scores. This annotation process adds valuable ground truth data for training and evaluating model performance (Fig. 1).

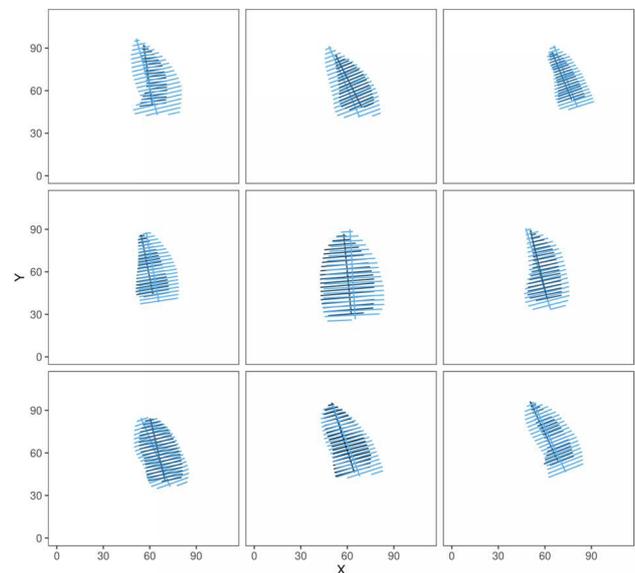


Fig. 1 EchoNet-Dynamic Annotation

### B. Model

1) *You Only Look Once*: The YOLOv8 model, a renowned algorithm for detecting objects, offers five distinct variations: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), and YOLOv8x (extra large). Each version can be further categorized into two specialized types, each serving a specific purpose in computer vision. One type is primarily focused on detection, capable of accurately identifying objects within images. Conversely, the other type is designed explicitly for segmentation, excelling in precisely delineating the boundaries and classifying the regions of the detected objects. It is important to note that the YOLOv8 model shares a similar backbone architecture with its predecessor, YOLOv5, based on a modified version of the Darknet architecture [23]. For the intentions of this study, the segmentation variant of YOLOv8x recognized as YOLOv8x-seg was employed.

2) *U-Net*: The U-Net structure has been intentionally designed to effectively tackle and overcome the diverse obstacles that arise when dealing with segmentation of medical images [24]. This model can be perceived as being split into two major components, specifically the contractive (encoding) path and the expansive (decoding) path. Since the input image traverses each block of the encoding path, it experiences a reduction in dimensionality by a factor of one-half. This particular dimensionality reduction allows the network to capture and consider the necessary contextual information effectively. Following this, during the decoding pathway, the image proceeds to undergo a process of up-sampling, allowing for a gradual and incremental restoration of its original dimensionality. It is during this expansive phase that the image is further refined, thus ensuring that the segmentation process is carried out in a highly detailed and accurate manner [25].

3) *Attention U-Net*: The Attention U-Net model is an advanced version of the traditional U-Net architecture, tailored explicitly for medical image segmentation tasks [26]. It integrates attention gates (AGs) to focus selectively on essential features within an image while suppressing irrelevant information, enhancing the model's accuracy and sensitivity to target structures. AGs are a crucial component of the Attention U-Net, enabling the model to concentrate on salient features useful for a specific task. These gates generate soft region proposals highlighting relevant features and suppressing the activations in irrelevant regions. The implementation of AGs does not require significant computational resources, nor does it substantially increase the model's complexity. The Attention U-Net model follows the standard U-Net architecture with an encoder-decoder structure but with AGs incorporated at each decoder level. These gates filter the feature maps coming through the skip connections from the encoder to the decoder, ensuring that only pertinent features are passed forward. This selective attention mechanism allows the model to maintain high sensitivity and specificity across various segmentation tasks [27].

4) *Annotation*: The EchoNet-Dynamic annotations were meticulously carried out by highly skilled and qualified professionals in the field of sonography and cardiology [13]. The annotations consist of two specific time points for each echocardiography. In Fig. 1, the pressure at end-diastole is represented by blue lines, whereas black lines delineate the pressure at end-systole.

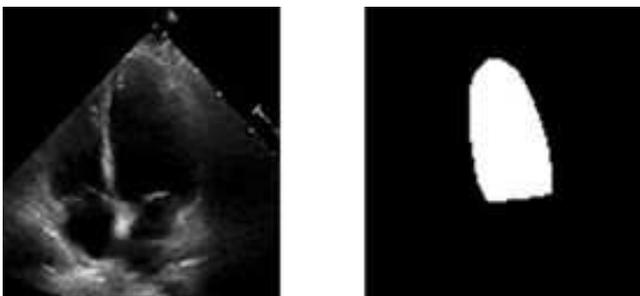


Fig. 2 EchoNet-Dynamic annotation converted to mask image

Given that the EchoNet-Dynamic annotation format significantly differs from the prevailing formats utilized by

YOLOv8 and U-Net, it was imperative to undergo a conversion process. Initially, the x and y coordinates encompassing the EchoNet-Dynamic annotation information were skillfully transformed into binary masking images, thereby facilitating the training of U-Net models (Fig. 1 and 2). Subsequently, the binary images were deftly reconverted back into their original x and y coordinates, all while ensuring an essential normalization step was incorporated.

5) *Model Training*: All of the models utilized in this research were subjected to training using the identical EchoNet-Dynamic dataset. To circumvent the potential problem of overfitting, the validation loss value and the Dice Coefficient Score (DCS) were actively monitored by implementing an early stop function for each model. During this investigation, it was decided that all models would undergo training on the EchoNet-Dynamic dataset. The DCS equation is shown in (1). This dataset was meticulously partitioned into distinct training, validation, and test sets, facilitating straightforward and unambiguous segregation between the various stages of model evaluation. The DCS was identified as the critical metric for validation, primarily due to its pertinence in evaluating the accuracy of segmentation. To prevent overfitting and ensure that the models retained their generalizability, an early stopping mechanism was meticulously established and subsequently employed throughout the model training process, with the basis for such a strategy being the DCS. Upon completing the training and validation phases, the models were meticulously assessed on an entirely separate test set, thereby allowing for the accurate quantification of their segmentation performance on previously unobserved data. This final step confirmed the models' ability to segment echocardiographic images effectively and precisely.

$$DCS (Ground\ truth, Predict) = \frac{2|Groundtruth \cap Predict|}{|Groundtruth| + |Predict|} \quad (1)$$

### III. RESULTS AND DISCUSSION

The evaluation of the models, conducted through the DCS to assess their segmentation accuracy, yielded significant findings. The YOLOv8 model, designed primarily for object detection, achieved a DCS of 0.8370. The U-Net model, tailored explicitly for image segmentation tasks, demonstrated superior performance with a DCS of 0.9236. Notably, the Attention U-Net model, an enhanced version of U-Net incorporating attention mechanisms, achieved the highest DCS of 0.9275. This incremental improvement suggests that the attention mechanisms within the Attention U-Net model contribute positively to segmentation accuracy.

Fig. 3 gives a schematic comparison of left ventricular segmentation on echocardiographic images and effectively presents the expert-designated ground truth versus the predicted contours by the YOLOv8 model. In the left-hand column, the boundary of the left ventricle (LV) is meticulously outlined by expert cardiologists, serving as the gold standard for evaluating both segmentation accuracy and the model's fidelity in replicating intricate cardiac structures. In contrast, the right column showcases the predictive capabilities of the YOLOv8 model, with algorithmically derived segmentation borders superimposed on the echocardiographic images. This direct visual comparison

provides an unimpeded, "at a glance" assessment of the model's alignment with the ground truth. The differences between the predicted segmentation and the ground truth pinpoint areas where the model excels and areas where improvement is needed, especially in the detailed detection and delineation of the left ventricle's finer contours.

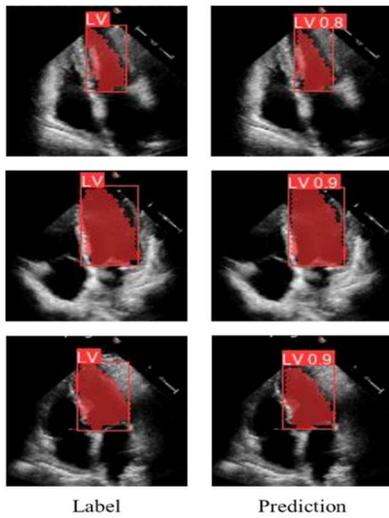


Fig. 3 Ground truth (Left) and predicted images (Right) of YOLO Model

In Fig. 4, adopting the same visual analytic approach as Fig. 3, the focus shifts to the segmentation performance of the U-Net model. This figure similarly displays the U-Net's predicted left ventricular segmentations alongside the established ground truth, with each row representing a case study. The U-Net model, known for its efficiency in deep learning for medical image segmentation, demonstrates its exceptional skill in capturing the complex geometry of the left ventricle. The side-by-side comparison highlights the U-Net model's capability to closely mirror the ground truth segmentation with remarkable precision, suggesting its potential to enhance diagnostic accuracy in echocardiography significantly.

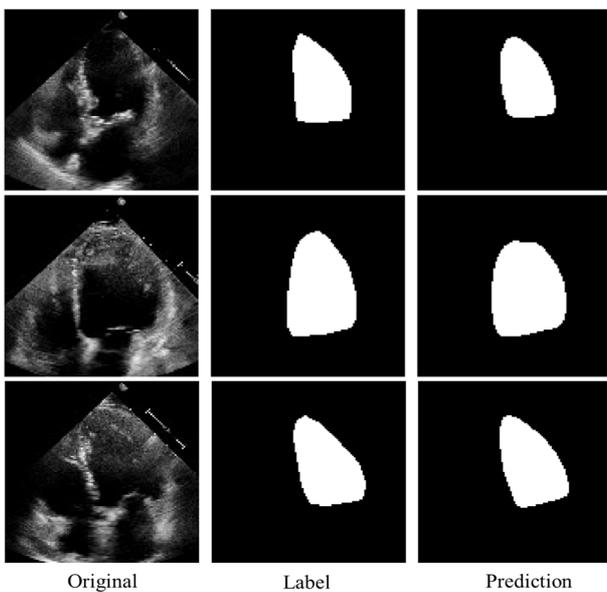


Fig. 4 Original images, ground truth masks, and predicted image of U-Net

Fig. 5 shows the results of the applied model Attention U-Net in echocardiographic images with the objective of cardiac structure segmentation. The images below are the original echocardiographic images in which the left ventricle could be seen in different views of the heart. The attention maps produced by the Attention U-Net model in Fig. 5 are from the right column and correspond to the respective example.

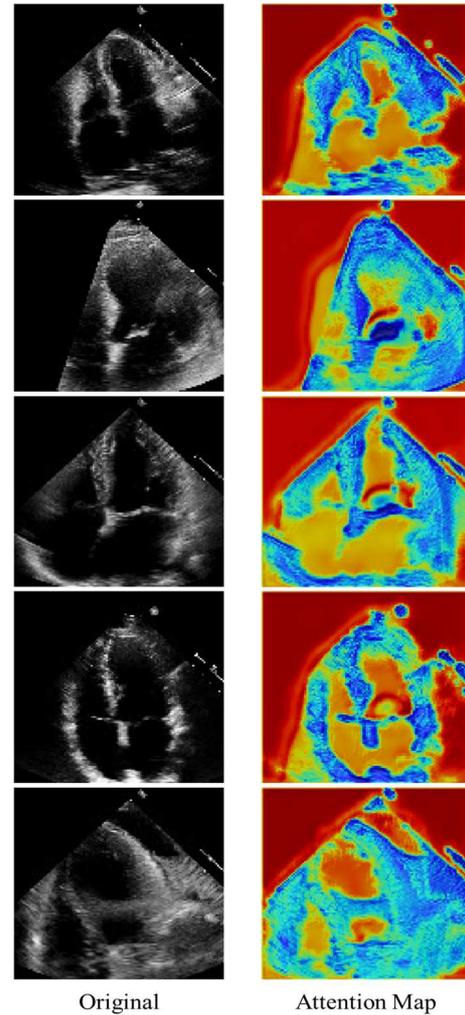


Fig. 5 Attention map generated from Attention U-Net

Regions in these maps where segmentation focuses most are colored red. The degree of redness on the map was an indicator of the degree of focus, which suggests more computational resources in some parts of the image that were judged to be of more importance in enabling the identification of the left ventricle. Attention maps help to understand how the Attention U-Net model interprets the way it processes the echocardiographic imagery to provide the final output and give insights into the model's decision-making. This visualization is used to validate whether the model learned well for left ventricle detection and provide a window into the patterns discovered by the model with respect to the cardiac anatomy in the images.

The results underscore the relative strengths of segmentation-focused architectures like U-Net and Attention U-Net over the more generalist YOLOv8 model in precise anatomical segmentation. The Superior performance of the U-Net architectures can be attributed to their design, which

optimizes for capturing spatial hierarchies and detailed contextual information essential for accurate segmentation of medical images [28]. While the YOLOv8 model's performance was commendable, its lower DCS indicates a limitation in handling the specific demands of medical image segmentation where precise delineation of boundaries is paramount. This may stem from the model's foundational design, prioritizing speed and object detection over the nuanced segmentation required for accurate medical analyses [29]. The limitations observed with YOLOv8 highlight the necessity for models that are adept at recognizing the presence of specific structures like the left ventricle and capable of accurately defining their contours.

The slight increase in DCS observed when comparing the Attention U-Net model to its U-Net counterpart can be attributed to the strategic inclusion of attention mechanisms within the former. These mechanisms optimize the allocation of computational resources to segments of the echocardiographic imagery that are most relevant for segmentation tasks. Unlike the conventional U-Net framework, which processes all regions of an input image indiscriminately, the Attention U-Net introduces attention gates. These gates implement a selective prioritization strategy, focusing on areas containing the left ventricle, thus improving segmentation accuracy [27], [30]. This focused approach is particularly beneficial when image clarity is compromised or when the left ventricle's boundaries are not clearly visible. Additionally, it enhances the model's ability to accurately define complex cardiac structures. Therefore, the modest improvement in DCS demonstrates the effectiveness of attention mechanisms in improving the specificity and efficiency of medical image segmentation, enabling a more targeted and effective identification of critical anatomical features.

This investigation into the utilization of advanced artificial intelligence (AI) models, specifically YOLOv8, U-Net, and Attention U-Net, for the segmentation of echocardiographic images has yielded significant insights into the capabilities and potential of AI to enhance cardiovascular diagnostics. The comparative analysis revealed that while YOLOv8 offers commendable object detection capabilities, the U-Net and Attention U-Net models demonstrate superior precision in segmenting the left ventricle, a critical task in cardiac function assessment.

The Attention U-Net model, in particular, showcased the highest segmentation accuracy, as evidenced by its Dice Coefficient Score (DCS). This increment in performance can be attributed to the model's incorporation of attention mechanisms, which allow for a focused analysis of relevant image features. Such mechanisms are crucial in medical imaging, where the precise delineation of anatomical structures is paramount. The results indicate a marked improvement in segmentation accuracy, pointing toward the effectiveness of specialized AI models in addressing the nuanced requirements of medical image analysis.

Moreover, the study highlights the transformative potential of AI in echocardiography, suggesting that adopting AI-driven models could significantly alleviate the operational strain on medical staff. By automating the segmentation process, these models promise to streamline clinical workflows, enhance diagnostic accuracy, and ensure a higher

standard of patient care. This research underscores the need for continued development and integration of AI technologies in clinical settings, particularly cardiovascular diagnostics.

#### IV. CONCLUSION

This study advances the domain of cardiovascular diagnostics through an in-depth examination of left ventricular segmentation within echocardiographic imagery. Utilizing sophisticated deep learning architectures, including YOLOv8, U-Net, and Attention U-Net, the study validates the substantial efficacy of these models in facilitating the automation of this crucial diagnostic process, achieving remarkable levels of accuracy. A thorough comparative analysis revealed that, among the evaluated models, Attention U-Net demonstrated superior performance, earning a Dice Coefficient Score of 0.9275. This outcome accentuates the critical role of attention mechanisms in augmenting the precision of image segmentation tasks. The implications of this research are profound, illustrating the transformative capacity of automated segmentation techniques to redefine the landscape of cardiac healthcare. The approach fosters a more comprehensive assessment of cardiac conditions by integrating advanced artificial intelligence models with diverse data modalities, including electrocardiogram (ECG) and magnetic resonance imaging (MRI). This integration bolsters diagnostic accuracy and optimizes clinical workflows, thereby enhancing patient care outcomes. The study highlights the distinct advantages of segmentation-focused architectures, particularly U-Net and Attention U-Net, over-generalized models such as YOLOv8 for tasks requiring detailed anatomical delineation. The employment of attention mechanisms within the Attention U-Net model is underscored as a pivotal strategy for maximizing computational efficiency and refining segmentation precision, especially under conditions of reduced image quality.

This research makes a notable contribution to the development of AI-driven diagnostic tools in cardiovascular care, opening new intelligent and efficient methodologies for evaluating and managing heart diseases. The present study wedes the frontiers of technology innovation with clinical expertise. It establishes the groundwork for prospective research projects designed to harness the potential of AI technologies in ways that may foster improved health delivery systems and patient outcomes.

#### ACKNOWLEDGMENT

This research was partly supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0023675, HRD Program for Industrial Innovation) and the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2024-RS-2022-00156439) supervised by the IITP (Institute of Information Communications Technology Planning Evaluation).

#### REFERENCES

- [1] C. A. Roberts, M. Binder, and D. Antoine, "Reflections on Cardiovascular Disease," *The Bioarchaeology of Cardiovascular Disease*, pp. 258–262, Apr. 2023, doi: 10.1017/9781108648561.019.

- [2] C. J. McAloon, F. Osman, P. Glennon, P. B. Lim, and S. A. Hayat, "Global Epidemiology and Incidence of Cardiovascular Disease," *Cardiovascular Diseases*, pp. 57–96, 2016, doi: 10.1016/b978-0-12-803312-8.00004-5.
- [3] A. Džubur, E. Begić, A. Durak-Nalbantić, and B. Aziri, "Cardiomyopathies," *Galenika Medical Journal*, vol. 2, no. 5, pp. 23–30, 2023, doi: 10.5937/galmed2305023d.
- [4] I. Kåks, M. Leopoulou, G. Mattsson, and P. Magnusson, "An Overview of the Cardiomyopathies," *Cardiomyopathy - Disease of the Heart Muscle*, Oct. 2021, doi: 10.5772/intechopen.95568.
- [5] K. A. Sytydykova, Y. V. Kazantseva, A. M. Sinanyan, M. N. Dulkina, G. O. Fatenkov, and M. A. Fatenkova, "Age-related risk factors affecting the population morbidity," *Cardiometry*, no. 27, pp. 221–228, May 2023, doi: 10.18137/cardiometry.2023.27.221228.
- [6] K. England and N. Azzopardi-Muscat, "Demographic trends and public health in Europe," *European Journal of Public Health*, vol. 27, no. suppl\_4, pp. 9–13, Oct. 2017, doi: 10.1093/eurpub/ckx159.
- [7] K. Fendrich, N. van den Berg, U. Siewert, and W. Hoffmann, "Demografischer Wandel," *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz*, vol. 53, no. 5, pp. 479–485, Apr. 2010, doi: 10.1007/s00103-010-1051-x.
- [8] R. M. Lang et al., "Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging," *European Heart Journal - Cardiovascular Imaging*, vol. 16, no. 3, pp. 233–271, Feb. 2015, doi:10.1093/ehjci/jev014.
- [9] H. Cui, L. Hu, and L. Chi, "Advances in Computer-Aided Medical Image Processing," *Applied Sciences*, vol. 13, no. 12, p. 7079, Jun. 2023, doi: 10.3390/app13127079.
- [10] B. S. Chandra, and S. Jana, "Towards Next-Generation Cardiac Care in Resource-Constrained and High-Risk Scenarios," Indian institute of technology Hyderabad, 2019.
- [11] H. Salah, "Improving cardiac care delivery using predictive and prescriptive analytics", University of Missouri–Columbia, doi:10.32469/10355/91687.
- [12] G. R. Djavanshir, X. Chen, and W. Yang, "A Review of Artificial Intelligence's Neural Networks (Deep Learning) Applications in Medical Diagnosis and Prediction," *IT Professional*, vol. 23, no. 3, pp. 58–62, May 2021, doi: 10.1109/mitp.2021.3073665.
- [13] D. Ouyang et al., "Video-based AI for beat-to-beat assessment of cardiac function," *Nature*, vol. 580, no. 7802, pp. 252–256, Mar. 2020, doi: 10.1038/s41586-020-2145-8.
- [14] S. Zheng, X. Ye, J. Tan, Y. Yang, and L. Li, "Dual-attention deep fusion network for multi-modal medical image segmentation," *Fourteenth International Conference on Graphics and Image Processing (ICGIP 2022)*, Jun. 2023, doi: 10.1117/12.2680056.
- [15] Y. Wang, C. Yin, and P. Zhang, "Multimodal Risk Prediction with Physiological Signals, Medical Images and Clinical Notes," *Heliyon*, May 2023, doi: 10.1101/2023.05.18.23290207.
- [16] T. K. M. Wang and A. L. Klein, "Multi-Modality Cardiac Imaging for Pericardial Diseases: A Contemporary Review," *Reviews in Cardiovascular Medicine*, vol. 23, no. 10, p. 336, Oct. 2022, doi: 10.31083/j.rcm2310336.
- [17] M. Wamil, M. Goncalves, A. Rutherford, A. Borlotti, and P. A. Pellikka, "Multi-modality cardiac imaging in the management of diabetic heart disease," *Frontiers in Cardiovascular Medicine*, vol. 9, Nov. 2022, doi: 10.3389/fcvm.2022.1043711.
- [18] N. Sriraam, S. TV, and S. Suresh, "A Computer Aided Markov Random Field Segmentation algorithm for assessing Fetal Ventricular Chambers," *Critical Reviews in Biomedical Engineering*, 2023, doi:10.1615/critrevbiomedeng.2023046829.
- [19] N. Joshi and S. Jain, "Improved Segmentation of Cardiac MRI Using Efficient Pre-Processing Techniques," *Journal of Information Technology Research*, vol. 15, no. 1, pp. 1–14, Aug. 2022, doi:10.4018/jitr.299932.
- [20] É. O. Rodrigues, F. F. C. de Morais, and A. Conci, "On the automated segmentation of epicardial and mediastinal cardiac adipose tissues using classification algorithms," *arXiv preprint arXiv:2208.14352*, 2022.
- [21] M. Penso et al., "Cardiovascular magnetic resonance images with susceptibility artifacts: artificial intelligence with spatial-attention for ventricular volumes and mass assessment," *Journal of Cardiovascular Magnetic Resonance*, vol. 24, no. 1, p. 62, Jan. 2022, doi:10.1186/s12968-022-00899-5.
- [22] S. P. Primakov et al., "Automated detection and segmentation of non-small cell lung cancer computed tomography images," *Nature Communications*, vol. 13, no. 1, Jun. 2022, doi: 10.1038/s41467-022-30841-3.
- [23] L. Glenn Jocher, Ayush Chaurasia. "Instance Segmentation," <https://docs.ultralytics.com/tasks/segment/>.
- [24] M. Bal-Ghaoui, M. H. El Yousfi Alaoui, A. Jilbab, and A. Bourouhou, "U-Net transfer learning backbones for lesions segmentation in breast ultrasound images," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 5, p. 5747, Oct. 2023, doi:10.11591/ijece.v13i5.pp5747-5754.
- [25] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4\_28.
- [26] A. M. Helmi Mahran, W. Hussein, and S. E. D. M. Saber, "Automatic Teeth Segmentation Using Attention U-Net," Jul. 2023, doi: 10.20944/preprints202306.1468.v2.
- [27] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, and B. Kainz, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.
- [28] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli, and D. Merhof, "Medical image segmentation review: The success of u-net," *arXiv preprint arXiv:2211.14830*, 2022.
- [29] J. Du, "Understanding of object detection based on CNN family and YOLO." p. 012029.
- [30] M. Melinščak, "Attention-based U-net: Joint Segmentation of Layers and Fluids from Retinal OCT Images," *2023 46th MIPRO ICT and Electronics Convention (MIPRO)*, May 2023, doi:10.23919/mipro57284.2023.10159914.