CS577 Project

# Project title

Fusion Modeling & Knowledge Distillation Optimization for Video-Based Cardiac Monitoring

融合建模与知识蒸馏优化在基于视频的心脏监测应用

# Team members

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Team Name | Rockets | | | |
|  | | | | |
| FirstName | Last Name | E-Mail | ID | Name(CN) |
| Wangshui | Kang | [wkang10@hawk.iit.edu](mailto:wkang10@hawk.iit.edu) | A20500653 | 康旺水 |
| Junhui | Zhan | [jzhan4@hawk.iit.edu](mailto:jzhan4@hawk.iit.edu) | A20499169 | 詹俊辉 |
| Sidi | Li | [sli178@hawk.iit.edu](mailto:sli178@hawk.iit.edu) | A20479715 | 李思迪 |

# Description of the problem

Despite significant advancements in the prevention and diagnosis of cardiovascular disease in recent years, heart disease remains the leading cause of adult mortality worldwide. This is partly attributed to the vital role played by blood circulation in facilitating oxygen-carbon dioxide exchange among most organs within the body. Consequently, any abnormality in the heart's pumping capacity can lead to irreversible damage across multiple organs within a short timeframe (3-8 minutes).

However, current human assessment of cardiac function primarily focuses on limited monitoring of cardiac beat cycles and specialized blood tests. The measurement of left ventricular ejection fraction (the ratio between changes in left ventricular end-systolic volume and left ventricular end-diastolic volume) stands as one of the most crucial indicators for evaluating cardiac function.但是传统的通过超声仪器对心脏左心室射血分数测量却不是金标准，也很难成为金标准。因为心超测量EF始终是半定量，测量的数值貌似是客观的，但进行测量的操作者却是主观的。临床过程中经常会发现，不同的操作者测量同一患者的EF，结果却大不相同，甚至是同一个操作者对同一患者进行EF测量，也会得到不同的结果。那么，临床诊疗过程中如何尽可能获取准确的EF以指导诊断和治疗方案呢？But the traditional measurement of the heart's left ventricular ejection fraction (LVEF) by ultrasound instrumentation is not the gold standard, and is hardly the gold standard. Because the measurement of EF by ultrasound is always semi-quantitative, the value of the measurement appears to be objective, but the operator who performs the measurement is subjective. In the clinical process, it is often found that different operators measure the EF of the same patient, but the results are very different, and even the same operator for the same patient EF measurement, but also get different results. So, how can we obtain as accurate an EF as possible to guide the diagnosis and treatment plan during clinical diagnosis and treatment?

# Brief survey

Discrepancies observed during ejection fraction assessments are partially due to common heart rate irregularities and computational challenges associated with manually tracking ventricular size for each beat. Additionally, physiological differences age, chest size, working and living environments, as well as behavioral habits contribute to variations in the basic contour of the heart even when examining individuals with normal physiological functions. Furthermore, different angles used during testing by non-cardiologists utilizing point-of-care ultrasound can significantly impact results.

Even after extensive training, physicians may still exhibit substantial divergent biases or omissions when assessing emergencies or rare diseases. Such discrepancies pose potential dangers alike. Therefore, there is an urgent need for rapid, efficient, cost-effective, accurate, reproducible, and quantifiable methods for assessing cardiac function.

# Proposed work

The acquisition of echocardiographic images is rapid, cost-effective, and free from ionizing radiation, making it the most widely utilized modality in cardiovascular imaging. To address this challenge, we propose "Fusion Modeling & Knowledge Distillation Optimization," which leverages a video-based deep learning algorithm called EchoNet-Dynamic. This approach combines feature extraction using multiple models (employing a window-based attention mechanism with multiple convolutional networks, a mixed model that enjoys the benefit of both self-Attention(maybe also Cross-Attention) and Convolution (ACmix), while having minimum computational overhead compared to the pure convolution or selfattention counterpart) and surpasses human experts in critical tasks such as left ventricle segmentation and ejection fraction estimation while significantly reducing computational requirements. Although Knowledge Distillation Optimization can optimize the neural network structure to speed up inference, it comes at the cost of decreased accuracy. Therefore, we hope to compensate for this loss by Attention modeling It is user-friendly, requires minimal or no parameter tuning, and can be executed efficiently on personal computers. Future prospects include extending its application to detect and monitor other organs, tissues, and body behaviors through video analysis to enhance auxiliary clinical diagnosis, treatment planning, and risk assessment. A longer term plan is to realize 3D reconstruction of the entire cardiac activity by means of ultrasound video from multiple angles to achieve a 360-degree view and avoid cross-sectional errors due to the acquisition angle of the ultrasound probe.

# Preliminary plan

|  |  |  |  |
| --- | --- | --- | --- |
| Milestone | Work content | Plan Date | Finish Date |
| 1. Setup Dev ENV | Hardware including GPU、PyTorch config ready | 2023/10/01 | 2023/09/25 |
| 1. Get Data | Register and download the dataset via the Stanford Artificial Intelligence in Medicine and Imaging (AIMI) Center Shared Datasets Portal | 2023/10/01 | 2023/09/25 |
| 1. Get Init code | Download EchoNet-Dynamic code from GitHub | 2023/10/01 | 2023/09/29 |
| 1. Run and record base test | Use the step 2 and 3 to get the code run and get expected results | 2023/10/15 |  |
| 1. Fusion Modeling | Use different method to build models to get better features | 2023/10/17 |  |
| 1. Run and record Modeling | Test to find out a good network architecture also the hyper-parameters | 2023/10/21 |  |
| 1. Intermediate Project Report | Introduction the problem、data、what have done so far、what remains to be done | 2023/10/27 |  |
| 1. Model Optimization | Use Knowledge Distillation to cutdown the network | 2023/11/10 |  |
| 1. Run and record Optimization | Test to find out a good network arch also the hyper-parameters | 2023/11/20 |  |
| 1. Analysis and summary for final project report | Summary of the problem, previous work, methods,  and results; the problem try to address methodology; observations from the experiments; Conclusions and future work | 2023/11/28 |  |
| 1. Final Project Presentation | 5-8 minutes; Describe the motivation and problem description; Briefly present the intuition behind the technical details (methodology); Algorithm and results (you can use a demo) | 2023/11/28 |  |

# Reference

1. Video-based AI for beat-to-beat assessment of cardiac function

Paper Published: 25 March 2020

<https://www.nature.com/articles/s41586-020-2145-8> (need to pay but we've already paid to download it)

1. EchoNet-Dynamic <https://echonet.github.io/dynamic/>
2. EchoNet-Dynamic Code: <https://github.com/echonet/dynamic>
3. EchoNet-Dynamic Data:7.04 GB, December 2020

Access the dataset via the Stanford Artificial Intelligence in Medicine and Imaging (AIMI) Center Shared Datasets Portal. Pls registor and follow the rules.

1. ACmix <https://github.com/LeapLabTHU/ACmix>
2. U-Net <https://arxiv.org/abs/1505.04597>
3. Deeplab <https://arxiv.org/abs/1412.7062>
4. On the Integration of Self-Attention and Convolution <https://arxiv.org/pdf/2111.14556.pdf>
5. <https://download.csdn.net/blog/column/12194563/129044015>
6. <https://zhuanlan.zhihu.com/p/539706748>
7. <https://zhuanlan.zhihu.com/p/539740657>
8. <https://zhuanlan.zhihu.com/p/608312950?utm_id=0&wd=&eqid=c37be367000f6cbf000000066486c9ce>
9. <https://stanfordaimi.azurewebsites.net/>
10. <https://www.ngui.cc/el/2485409.html?action=onClick>
11. <https://blog.csdn.net/qq_40280673/article/details/127449624>
12. <https://zhuanlan.zhihu.com/p/640904026>
13. <https://aisle.hzau.edu.cn/info/1097/1412.htm>
14. <https://blog.csdn.net/zuzhiang/article/details/107418459>

# Others

In this paper, we present the basic methodology of this computational framework, show the results of an extensive validation study, and demonstrate its utility as a functional metric.在本文中，我们介绍了这一计算框架的基本方法，展示了大量验证研究的结果，并证明了作为功能指标的实用性。

Thus, we present “MicroBundleCompute,” a computational framework for automatic quantification of morphology-based mechanical metrics from movies of cardiac microbundles. Briefly, this computational framework offers tools for automatic tissue segmentation, tracking, and analysis of brightfield and phase contrast movies of beating cardiac microbundles. It is straightforward to implement,requires little to no parameter tuning, and runs quickly on a personal computer. In this paper, we describe the methods underlying this computational framework, show the results of our extensive validation studies, and demonstrate the utility of exploring heterogeneous tissue deformations and strains as functional metrics.

With this manuscript, we disseminate “MicroBundleCompute” as an open-source computational tool with the aim of making automated quantitative analysis of beating

cardiac microbundles more accessible to the community.

However, there are limited accurate, reliable, and reproducible computational metrics for making quantitative comparisons of functional behavior.

streamlined, This underscores

the need for streamlined, accurate, and high-performance computational

tools.

However, when it comes to making quantitative comparisons of functional behavior, there are limited options for reliably and reproducibly computing functional metrics that are suitable for direct cross-system comparison. In addition,

the current standard functional metrics obtained from time-lapse images of cardiac microbundle contraction reported in the field (i.e., post forces, average tissue stress) do not take full advantage of the available information present in these data (i.e., full-field tissue displacements and strains).

<https://www.bilibili.com/video/BV1RY4y1P7uL/?spm_id_from=333.337.search-card.all.click&vd_source=a40c1f315886f97ba7018d224724764d>

<https://zhuanlan.zhihu.com/p/58761927>

<https://blog.51cto.com/u_15298598/6274394>

<https://blog.csdn.net/m0_75272311/article/details/130302448>

<https://www.rstk.cn/news/141859.html?action=onClick>

<http://wed.xjx100.cn/news/125890.html?action=onClick>

https://rrc.cvc.uab.es/?ch=8&com=download

## deeplab

<https://zhuanlan.zhihu.com/p/385299424>

<https://betheme.net/qianduan/59619.html>

<https://www.ngui.cc/el/2868859.html?action=onClick>

## Description of the problem

A brief survey of what have been done and how the proposed work is different.

Preliminary plan (milestones) and Reference (a list of papers)

## Reference：

This download URL can’t shere

https://aimistanforddatasets01.blob.core.windows.net/echonetdynamic-2/EchoNet-Dynamic.zip?sv=2019-02-02&sr=b&sig=khf02%2FaYvtW7yG928t04gC4U5maRWL4KMPsWoQFjBMg%3D&st=2023-09-25T14%3A45%3A56Z&se=2023-10-25T14%3A50%3A56Z&sp=r

<https://arxiv.org/search/?searchtype=all&query=cardiac+function&abstracts=show&size=50&order=-announced_date_first>

<https://arxiv.org/search/?searchtype=all&query=deeplab&abstracts=show&size=50&order=announced_date_first>

1. <https://pytorch.org/audio/stable/tutorials/speech_recognition_pipeline_tutorial.html>
2. <https://pytorch.org/tutorials/intermediate/speech_command_classification_with_torchaudio_tutorial.html>

https://www.bilibili.com/video/BV1Se411P7mw/?spm\_id\_from=333.337.search-card.all.click&vd\_source=a40c1f315886f97ba7018d224724764d 热门的cv都发展到3D了 CVPR 2022 Tutorial on Neural Fields in Computer Vision

Datasets：

1. <https://www.bilibili.com/video/BV1LR4y1x7Pw/?spm_id_from=333.337.search-card.all.click&vd_source=a40c1f315886f97ba7018d224724764d>

<https://www.bilibili.com/video/BV1hM4y157xX?p=11&vd_source=a40c1f315886f97ba7018d224724764d>

<https://www.bilibili.com/video/BV1TD4y137mP?p=20&vd_source=a40c1f315886f97ba7018d224724764d>

<https://paperswithcode.com/task/speaker-diarization>