Real-time AI-empowered echocardiography for Intensive Care Units in low- and middle-income countries

Miguel Xochicale *et al.*

July 21, 2022

Contents

1 Introduction **2**
1.1 AI-empowered echocardiography . . . 2
1.2 Real-time AI-empowwered echocardiography . . . 4

2 Methods for AI-empowered echocardiopgraphy **4**
2.1 Image Quality Assessment . . . 4
2.2 Classification of echochardiograms . . . 4
2.2.1 Classification of other US images . . . 5
2.3 Clustering techniques . . . 5
2.4 Auto-encoders . . . 5
2.5 Segmentation . . . 6
2.6 Contrastive Learning . . . 6
2.7 AI-guided US imaging . . . 6
2.8 3D US . . . 7
2.9 Transformers . . . 7
2.10 Others . . . 8

3 Spatiotemporal Features **8**
3.1 Deep Residual Learning . . . 8
3.2 LSTM . . . 8

4 Validation and Usability Studies **8**

5 Cardiac Motion Estimation **9**

6 Datasets in echocardiography **9**
6.1 CAMUS (2D US) . . . 9
6.2 EchoNet-Dynamic (2D US) . . . 9
6.3 CETUS (3D US) . . . 9
6.4 Synthetic cardiac motion . . . 10

7 Tools **10**
7.1 Annotation tools . . . 10

8 Methods and materials **10**
8.1 Proposed model 10

9 Datasets **10**
9.1 VITAL 10
9.2 Ethics statement 10

10 Potential future work **11**

1 Introduction

In the last decades the use of echocardiography is an important clinical approach in Intensive Care Units (ICU) because of the adtages of US devices such as portability, low cost, low radiation and its real-time capabilities to access cardiac anatomy [Feigenbaum, 1996, Vieillard-Baron et al., 2008, Singh and Goyal, 2007, Campbell et al., 2018]. Despite the previous advances, there various challenges in the current practices of clinical ultrasound:

- Intra-view variability of echochardiograms (physiological variations of subjects and adquisition parameters) and sonographer expertise, [Khamis et al., 2017, Feigenbaum, 1996, Field et al., 2011].
- Inter-view similarity of echochardiograms (similar views of valve motion, wall motion, left ventricle, etc) and transducer position during adquision, [Zhang et al., 2018], redudant information in the clinical echo system (icons, date, frame rate, etc). [Khamis et al., 2017]
- Limited number of expert clinitians to perform US imaging analysis and provice accurate diagnosis, and equipment and hospitalisation requirements in low- and middle-income countries (LMICs) [Hao et al., 2021].

1.1 AI-empowered echocardiography

[Hanson III and Marshall, 2001] reviwed various applications of AI in the ICU where real-time analaysis of waveforms of electrocardiograms and electroencephalograms using neuoral network were used to idenfy cardiac ischemia and diagnosis myocardial ischemia. [Hanson III and Marshall, 2001] also reviewed various scnerations where AI is used in the ICU, such as bayesian networks considering central venous pressure (CVP), left ventricular ejection fraction (EF), heart rate (HR), hemoglobin (HGB) and oxygen saturation (O2sat) resulting in a probablitistic cardiac output. [Hanson III and Marshall, 2001] also touched on datavisualiastion to demonstrate the hypothetical ICU for large number of patients (head injury, sepsis, acute respiratory distress syndrome, etc).

Ghorbani et al. in 2020 reported the first deep learning model to predict age, sex, weight and height from echocardiogram images and make use of such models to understand how models predicts systematic phenotypes which are difficul for human interpreters [Ghorbani et al., 2020]. Authors trained CCN models with 2.6 million echocardiogram images from 2850 patients with the extraction of labels local strucutre and features (e.g. pacemaker lead, dilation of left atrium, hypertrophy for left ventricular) and labels from the physician-interpreted report (e.g, catheters, pacemaker, and defibrillator leads). Recently, [Hong et al., 2022] reviewed 673 papers that made use of machine learning-enabled to help for clinical desicion in the ICU, of these studies the majority used supervised learning (91%) few doing unsupervising learning and reinforcement learning. Similarly, [Hong et al., 2022] identified 20 of the most fequent variables in machine learning-enabled in the ICU, being the top five (age, sex, heart rate, respiratory rate, and pH). [Hong et al., 2022] mentioned that typical outcomes in the ICU are mortality, survival, and long-term quality of life and included typical patient outcomes, specific diseases, and stay of time evaluation.

Tromp et al. classified a dataset of 1145 2D echocardiography videos as apical 4 chamber (A4C) view, apical 2 chamber (A2C) view, parasternal long axis (PLAX) view, or 2D other views and focused versions of the main views [Tromp et al., 2022]. Authors used CNN of four layers, dense network and softmax output layer, trained with categorical cross-entropy loss function, then a second classifier of an unsupervised deep learning clustering CNN, trained with mean square error and Kullback-Leibler loss functions [Tromp et al., 2022].

[Zhang et al., 2022] reviewed AI's applications in left ventricular systolic funciton (LVEF) and global longitudinal strain (GLS), pointing out its dependendant to the sonographers's expertise (inter-observer variability) and post-processing and variablity in diferen US devices. [Zhang et al., 2022] pointed the challenges of AI-enhanceed echocargiografy for interpretability of results and its sensitivity to sample shortage, to which authors mention about the potentials of multimodal imaging (us, mri and clinical data) to improve detection rate of diseases.

[Behnami et al., 2020] applied DenseNet-like network for feature learning and RNN unit with bidirectional Gated Recurrent Units to alivate loss of information from the ealier frames of echos to automatically detect high risk or low risk of heart failure with reduced ejection fraction with an overall accuracy of 83.15%, precision of 82.6% and recall of 81.1%. [Behnami et al., 2020] mentioned that EF is highly user-dependant to which they propose to collect more data,

[Liu et al., 2021] proposed pyramid local attention neural network (PLANet) to improve segmentation performance of automatic methods in 2D echocardiography. PLANet was evaluated with CAMUS and sub-EchoNet-Dynamic datasets, showing a better performance against geometric and clinical metrics.

[Ulloa Cerna et al., 2021] made use of DNN to learn spatiotemporal features from echocardiography video data to enhance clinical prediction of 1 yr all-cause mortality where video echo data linked to EHR data that included hand-crafted echocardiography-derived measurements (EDMs), additional clinical variables and individual outcomes. The DNN model presents "superior prediction performance" over four cardiologist and two benchmark clinical models: the pooled cohort equations (PCE) and Seattle Heart Failure (SHF) risk score [Ulloa Cerna et al., 2021]. [Ulloa Cerna et al., 2021] used "full, raw (annotation-free) echocardiographic videos to make predictions by learning from more than 812,278 clinically acquired echocardiography videos of the heart (50 million images)."

[Jafari et al., 2021] pointed out the challenges of obtained high quality for less experience operators and the hight variablity or echo quality adn cardiovascular structures accross different patients to which authors proposed "Bayesian deep learning approach for fully automatic LVEF estimation based on segmentation of the left ventricle (LV) in parasternal short-axis papillary muscles (PSAX-PM) level". [Jafari et al., 2021] made use of 2,680 patients with PSAX-PM echo cine acquired by a variety of ultrasound devices, namely iE33, Vivid i/7/9/95, Sonosite, and Sequoia (only 554 echo cines were considered as ground truth with LV mask delineated by an experienced level III echocardiographer).

Author(s), year	Datasets (clips)	Input frames	Flops	Model; Params	Latency	Hyper-parameters	Code
[Wu et al., 2022]	EchoDynamic (10Kc) CAMUS (.5Kc)	3	56.359	TCE-BSSF;74.798	32 (ms/f)	lr=10e-4;epochs=100; CR=5-fold	-
[Van Woudenberg et al., 2018]	AP4(2Kc)	10	-	DenseNet-LSTM;	352.91ms	-	- -
[Østvik et al., 2021]	Synthetic/Simulated/Clinical	7	-	Echo-PWC-Net	-	-	-

Ono et al. applied differen models where Unet++ demonstrated good performance for automatic endocardial border detection and left ventrical assesment in 2D echocardiographu videos [Ono et al., 2022]. The datasets to train networks was made of 2798 images from 118 videos of which 22 videos with 465 frames were for 4CV [Ono et al., 2022]. Ono et al. also touched on the challenges of providing explainable AI for US imaging.

1.2 Real-time AI-empowwered echocardiography

In terms of real-time analysys of echocardigraphy, , Wu et al. 2022 applied baselines of UNET with temporal context-aware encoder (TCE) and bidirectional spatiotemporal semantics fusion (BSSF) modules to EchoDynamic (10030 video sequences with of 200frames of 112x112 pixes) and CAMUS datasets (450 video of 20 frames of 778x594 pixels) with evaluation metrics of Dice score (DS), Hausdorff Distance (HD), and area under the curve (AUC) [Wu et al., 2022]. Wu et al. 2022, ensuring low latency and real-time performance, presented speed analysis againts eight methods using calculations number FLOPS (G), number of parameters (M) and speed (ms/f) which lowest one was 32 ms/f [Wu et al., 2022].

2 Methods for AI-empowered echocardiopgraphy

2.1 Image Quality Assessment

[Labs et al., 2021] considers chamber clarity, depth gain, on-axis attributes, apical foreshoredness.

2.2 Classification of echochardiograms

[Van Woudenberg et al., 2018] applied DenseNet and LSTM to extract temporal information on sequences of 16K echo cine frames to classify 14 heart views with an average accuracy of 92.35%. [Van Woudenberg et al., 2018] implemented a Tensorflow runner that performs contrast enhancement to then sent each frame to three identical CNNs running in separated threads to prevent lag during inference times. Then a shared buffer collects extracted features from CNNs to then awake the thread for the LSTM network from the previous ten frames to produce classification and quality prediction. [Van Woudenberg et al., 2018] also presents timing diagrams to quantify frame arrival and real-time performance to operate at 30 frames per second, while providing feedback with a mean latency of 352.91 ± 38.27 ms when measured from the middle of the ten-frame sequence.

Zhang et al. performed view classification with 277 echochardiograms to create a 23-class models (including a4c no occlusions, a4c ocluded LA, a4c ocluded LV, etc) using 13-layer CNN with 5-fold cross-validation for accuracy assement and resulting in 84% for overall accuracy where challenges for partial obscured LVs for a2c, a3c and a4c [Zhang et al., 2018]. Similarly, Zhang et al. applied U-net to segment 5 views (a2c, a3c, a4c, PSAX, PLAX) and CNN model for 3 chardiac disseases with the use of A4c captuting most of the information for the diseases.

Khamis et al. considered 309 clinical echocardiogram of apical viewss which were visiually classified and labelled by two experts into three classes: 103 a2c views, 103 a4c views and 103 alx views to then applied spatio-temporal feature extraction (Cuboic Detector) and supervised learning dictionary (LC-KSVD) resulting in an overall recognition rate of 95% [Khamis et al., 2017].

2.2.1 Classification of other US images

[Baumgartner et al., 2017] proposed SonoNet which is a VGG-based arquitecture, SonoNet64 used the same first 13 layers of VGG16, and SmallNet, loosely inspired by AlexNet, for real-time detection and bounding box localisation of of standard views in freehand fetal US. [Toussaint et al., 2018] applied four feature extraction networks couple with batch-normalisation and soft proposal layer (VGG13-SP, VGG16-SP, ResNet18-SP, ResNet34-SP) being ResNet18-SP the best performing network with avarage accuracy over six classes of fetal US views (0.912). Authors mentions that detection and localisation of anatomical views were tested in real-time performance at inference time (40ms per image, or 20Hz). [Xie et al., 2020] proposed a dual-sampling convolutional neural network (DSCNN) for US image breast cancer classification, being DSCNN more efficent than AlexNet, VGG16, ResNet18, GoogleNet and EfficientNet. [Snider et al., 2022] reported summaries of CNN heuristics to detect shrapnel in US images. Authors presented summaries of model performance for layer activators, 2D CNN layer arquitectures, model optimisers dense nodes, and the effect of image augmentation and dropout rate and epoch number. [Boice et al., 2022] proposed ShrapML, a CNN model to detect shrapnel in US imaging. Authors compared ShrapML (8layers–6CNN,2FC, 0.43 million of parameters) against DarkNet19, GoogleNet, MobileNetv2 and SqueezeNet, being ShrapML 10x faster than MobileNet2 which offered the highest accuracy.

2.3 Clustering techniques

Zhang et al. mentioned that 23 view classes from 7168 individually labeled videos that ware classified with a 13-layer CNN to then viewed with the use of t-Distributed Stochastic Neighbor Embedding [Zhang et al., 2018]. Zhang et al. made use of 277 echochardiograms collected over a 10-year period for view classification. Kusunose et al. mentioned that other authors have reached an acciracy of 91-94 for 15-view classification while their work mentioned a 98.1 accuracy for five-prederminted views [Kusunose, 2021].

2.4 Auto-encoders

Laumer et al. proposed a novel autoencoder-based framework to learn human interpretable representation of cardiac cycles from cardiac ultrasound data [Laumer et al., 2020],

Ouyang et al. presented echo-dynamic dataset as the first annotated medical video dataset with 10,036 videos. Additionally, authors reported the use of three CNN arquitectures varing filters in each layer to assess ejection fraction to near-expert performance. It is worthwhile to note that authors got best performance with mean absolute error of 5.44% using clip lenght of 16 and frame rate of 4. Such error is near-expert perfonace as they can get 4-5% for skilled echochardiographers in cotrolled settings [Ouyang et al., 2019].

Ghorbani et al. applied convolutional neural networks of cardiac ultrasound to identify local structures, estimate cardiac function and predict pathologies. Their deep learning model, EchoNet, can identify up to 10 cardiac biometrics which results in decreasing repetitive task in the clinical flow, provide interpretation to less experienced cardiologist, and predict phenotipes. This work can predict age, sex, weight and heigth from echocardiogram images. Authors mention that the increase of date does not improve model traiining. The homogenisation of cadiac views prior to model training improved training speed and computaitonla time [Ghorbani et al., 2020]

2.5 Segmentation

With the challenges of limited sampling of cardiac cycles and the considerable inter-observer variability, Ouyang et al. presented a CNN model with residual connections and spatiotemporal convolutions that surpase human performance of segmentaion of left ventricle, estimation of ejection fraction and assessment of cardiomyophaty. Their model reached Dice similarity coefficient of 0.92, predicts ejection fraction with mean absolute error of 4.1% and clasify heart failure based on reduced ejection fraction [Ouyang et al., 2020].

Meyer et al. used Prominence Iterative Dijkstra's algorithm (ProID), based on the identification of ventricle boundaries with iterative Dijkstra's algorithm, for ventricle detection and volumen estimation [Meyers et al., 2021]. ProID employs echochardiogram-specific cost-matrix to address contrast-to-noise and resolution limitations problems [Brindise et al., 2020

2.6 Contrastive Learning

Methods on Contrastive Learning apparently address the challenge of required labelled data to identify pathologies in the images of dectect certain cardiac views. Recently, Chartsias et al. use contrastive learning to train imbalanced cardiac datasets and they compared a naive baseline model to achieve a F1 score of up to 26% [Chartsias et al., 2021] Saeed et al. recently investigated contrastive pretraining to improve the DeepLabV3 and UNET segmentation networks of cardiac structers in ultrasound imaging. Authors showed comparable results with state-of-the-art fully supervised algorithms and presents better results compared to EchoNet-Dynamic and CAMUS [Mohamed et al., 2021]

2.7 AI-guided US imaging

Near-human quantification of LV and EF has been investated, however Asch et al. pointed out that buoundary identification is prone to errors when low quality images or artifacts are used Asch et al. pointed out that data and materials were not publicly available and they made use of AutoEF by captionhealth co. Authors used a databes of 50000 echocardiography datasets over a period of 10 years of varios clinical US syustems. The training datasets included multiople views of 2 and 4-chamber views and LV EF values where clininias use conventional methods (biplane Simpson technique) [Asch et al., 2019].

Asch et al. [Asch et al., 2021].

Hong et al. reported the evalition of imagin quality asssement to demostrated that AI can recognise nuaces of varing imaing during scanning [Hong et al., 2021]

Narang et al. reported the adquisiton of 10 echocardiography views of novices users using deep-learning-based software [Narang et al., 2021]. Narang et al. mentioned that CNN were used with stacks of networks and transofrmations. The AI-guided software consist of three estimates: (1) quality image assement, (2) "6-dimensional geometric distance with postion and orientation between the current probe location and the locattion anticipated to optimise the image"; and (3) corrective probe manipulation. [Narang et al., 2021] Authors mention that algorithms do not use trackers, fiducial marks or additional sensors to made guide estimations [Narang et al., 2021].

Cheema et al. reported the use of AI-enabled guidance to sonoographer which was created from the use of 500000 hand movmentes. Cheema et al. reported that such feature was the first cardiac aotorhisedd by Food and Drug adminstation in 2020. Authors presented five cases covid-19 intensive care unit (ICU) to illustrate "how desition making affect in patient care" and how the use of AI-enabled provided real-time guidance to acquire desired cardiac UL with the sterting of user's transducer position and hand movevemnt [Cheema et al., 2021].

2.8 3D US

Considering that 3D left ventricle (LV) can provide full volume information of the hearth than 2D echordaiography, Dong et al. proposed a real-time framework VoxelAtlasGAN that made use of cGAN [Dong et al., 2018]. VoxelAtlasGAN framework with mean surface distance of 1.85 mm, mean hausdorff distance of 7.66mm, mean dice 0.953 and correlation of EF 0.918 and the mean inference speed of 0.1 s demostrated potential for clinical applicaiton [Dong et al., 2018]. Dong et al. in 2020 applied tranformers to obtain translations paremeters that passed to VoxelAtlasGAN [Dong et al., 2020]. AlasNET framework ended up with "mean surface distance, mean hausdorff surface distance, and mean dice index were 1.52 mm, 5.6 mm and 0.97 respectively" [Dong et al., 2020]

Smistad et al. 2021 made use of CETUS 3D US LV segmentation dataset and weakly annorated datasets for real-time 3D left ventricle segmentation and estiomation of ejection fraction [Smistad et al., 2021b]. Authors presented the impact of pre-training that resulted in an improvement of Dice score. It is imporant to note that VoxelAlasGAN and AtlasNet by Dong el al. presented a better dice score. Smistad et al. 2021 concluded that a limited labelled datasets of 15 patiens demostrate good accuracy and models were able to generalise to new data and ultrasoudn scanners [Smistad et al., 2021b].

2.9 Transformers

Rubin et al. noted the shortcoming of transformers of extensive compation for training that lead to use detection transformer (DETR) which make smaller models reducing model size and acceleration inference [Rubin et al., 2021]., Rubin et al. considered the detection of needles in real-time ultrasound video sequences 12,000 needle insertions (2 million of individual frames). Video sequences (up to 60 sec in time) were divided into 30-frame clips (1 sec in time).

Reynaud et al. 2021 adapted Residual Autoencoder Network and BERT model to predict ejection fraction which is different from what is commonly use with segmentation methods [Reynaud et al., 2021]. Reynaud et al. applied their model to Echonet-Dynamic dataset which only constains 10,030 echocardiograms containing one to three or more cadiac cycles with only cardiac cycle with ES and ED annorations. Due to the distribution between ES and ED, the sequence legnth was 128 frames. As Echonet-Dynamic datasets contains unlabelled ES and ED, Reynaud et al. applied (a) Guided Random Sampling (b) Mirroring Methods. Code is available at `https://github.com/HReynaud/UVT`.

2.10 Others

Rank-2 non-negative matrix factorization [Yuan et al., 2017] to generate End-Systole and End-Diastole for apical 4 view. Recently Robust Non-negative Matrix Factorization seems to be implement low-computation cost algorithms to automatic segment mitral valve [Dukler et al., 2018].

Salte et al. classified three standard apical views from data of 200 patients to peformf straing measurements with deep learning arquitectures [Salte et al., 2021]. Salte el al. made use of the work [Østvik et al., 2021] inception and dense network were used to clasity, recurrent network to detect event timing and u-net-based network for segmentation [Salte et al., 2021]. Authors compared the results with the commenrially avaible semiautomatic speckle-tracking software (EchoPAC v202), reporting evinde of the comparable GLS measurmentes to other semiatuomatic methods [Salte et al., 2021].

3 Spatiotemporal Features

3.1 Deep Residual Learning

Ouyang et. al. benmarked various spatiotemporal convolutions (Sports-1M, Kinetics, UCF101, and HMDB51) [Ouyang et al., 2019] based on deep residual learning (He et al. 2015 and Tran et al. 2018).

3.2 LSTM

Recently, Smistad et al. 2021 presented the use of LSTM to address the single frame segmentation of end-diastole and end-systole to address segmentation fricklering and reduce tmeporal errors [Smistad et al., 2021a]. One of the challenges is architecture design to add ConvLSTMs to whcih authors experiment at the location at the endocer, decoder, last layer and in bottlenet, to which authors mentiosn that the use of the ConvLSTM layers in the encoder of the temporal NN gave the best results [Smistad et al., 2021a]. Authors mention that interpolation of the annorations of the entire cardiac cycle did not captured the complex motion witht he use of 7 frames to which they suggest to use advance speckle tracking such as Echo-PWC-Net [Østvik et al., 2021].

Lu et al. made use of U-Net and LSTM to model Left Ventricular cardiac motion [Lu et al., 2020].

Bar et al. assessed surgical workflow recongnition and report a deep learning system in wich clips of one second to create a short-term spatio-temportal model based on inflated 3D network with non-local to optain SoftMax probability vector to fed LSTM to produce final phase predictions [Bar et al., 2020]

4 Validation and Usability Studies

Francesconi et al. 2021 presented technical vaidation by intra-operator reproducbility of two measuremnts (intima-media thickness and distension) by one expert and usability of integrated system in a laboratory setting with 12 healthy volunteers [Francesconi et al., 2021].

5 Cardiac Motion Estimation

Cardiac motion estimation has important role in echocardiography due to the computation of myocardial deformation indices. [Aviles et al., 2017] propose a method to estimate cardiac motion of ultrafast ultrasound data by combining low-rank data representation with topology preservation to overcome the challenges of non-rigid registration that involves complex heart motion and distortions. [Young et al., 2021] proposed a method to build 3D recovery models from 2D echocardiography data with the only assumption that the probe is fixed for one cardiac cycle. [Evain et al., 2022] proposed a deep learning model PWC-Net, achieving an average endpoint error of ”0.07$\pm$ 0.06 mm per frame and 1.20$\pm$0.67 mm between ED and ES on their simulated dataset”. See [Østvik et al., 2021] for further details on Echo-PWC-Net and other networks.

6 Datasets in echocardiography

6.1 CAMUS (2D US)

CAMUS dataset, Cardiac Acquisitions for Multi-structure Ultrasound Segmentation, was publised in 2019 by Leclerc et al. 2019 [Leclerc et al., 2019]. CAMUS is the largest publicy-available and fully-annotated dataset of two and four-chamber adquistion from 500 patients. Datasets is cathegorised in image quality (good, medium, and poor) and LV_{EF} ($\leq$ 45% (phatological risk) , $\geq$ 55%, else). The dataset reflects a daily clinical practice data where images quality and a range of phatological cases. Dataset was collected with GE Vivid E95 ultrasound scanners (GE Vingmed Ultrasound, Horten Norway) with a GE M5S probe (GE Healthcare, US). The datasets is available electronically to download at `https://www.creatis.insa-lyon.fr/Challenge/camus/`.

6.2 EchoNet-Dynamic (2D US)

Ouyang et al. published a large datasets of 10,030 annotated echocardiogram videos [Ouyang et al., 2019, Ouyang et al., 2020]. Datasets were labelled left ventricle volumes by sonographers to calculate ejection fraction. Datasets were acquired by skilled sonographers using iE33, Sonos, Acuson SC2000, Epiq 5G or Epiq 7C ultrasound machines and processed images were stored in a Philips Xcelera system. The datasets is available electronically to download at `https://echonet.github.io/dynamic/index.html#dataset`.

6.3 CETUS (3D US)

Challenge on Endocardial Three-dimensional Ultrasound Segmentation (CETUS) dataset was published in 2016 by Bernard et al. [Bernard et al., 2016]. CETUS contains 45 sequences of 3D ultrasound volumes of one cardiac cycle from 45 patients were equally acquired from three different hospitals with three different brands of ultrasound machines (GE, Philips and Siemens) [Bernard et al., 2016]. The studied population of 45 participants is composed of 15 healthy subjects, 15 with previous myocardial infartation, 15 with dilated cardiography. The datasets is available electronically to download at `https://www.creatis.insa-lyon.fr/EvaluationPlatform/CETUS/about_database.html`.

6.4 Synthetic cardiac motion

Alessandrini et al. in 2018 published "a open access library of 105 synthetic sequences encompassing i) healthy and ischemic motion patterns, ii) most common apical probe orientations iii) vendor specific image quality from 7 different systems" [Alessandrini et al., 2018]. See previuos work of Alessandrini et al. in 2015 [Alessandrini et al., 2015].

7 Tools

7.1 Annotation tools

Recently, Smistad et al. 2021 published the first web-based tool for annoration of medical ultrasound video to do image classification, segmentation, bouding box and landmark annotation [Smistad et al., 2021c]. AW tool has been used since 2016 at different projects to perform segmentation of the left ventricle, cardiac view classification, and detection of nerves and blood vesels [Smistad et al., 2021c].

8 Methods and materials

8.1 Proposed model

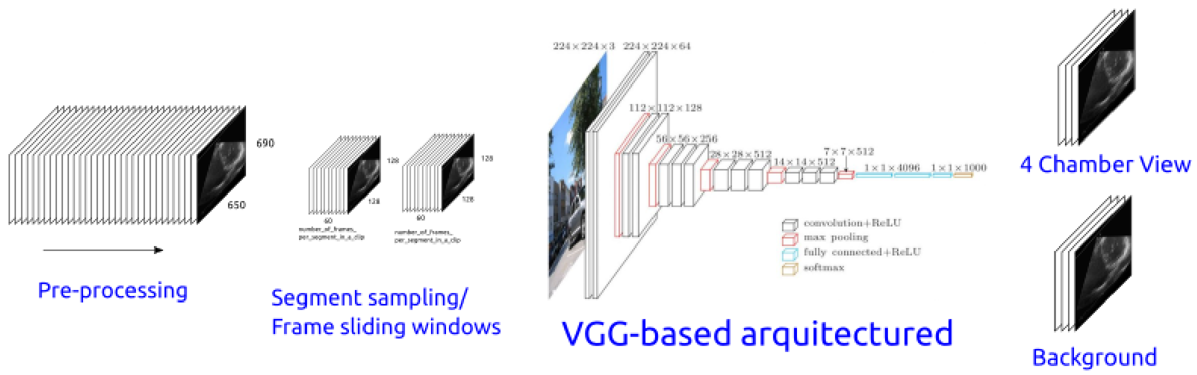

Figure 1: VGG-based arquitecture (**a**) description... (**b**) description... Figure is adapted from the works of

9 Datasets

9.1 VITAL

86 patients of average age (?) ? male and ? female were collected by four clinicians of ? years of expience collected echochardiography datasets. The collection was done with the clinical device GE Venue Go machine and GE convex probe C1-5-D.

9.2 Ethics statement

This study was approved by ... and the ethics committee ... All participants gave written informed consent to participate before enrollment.

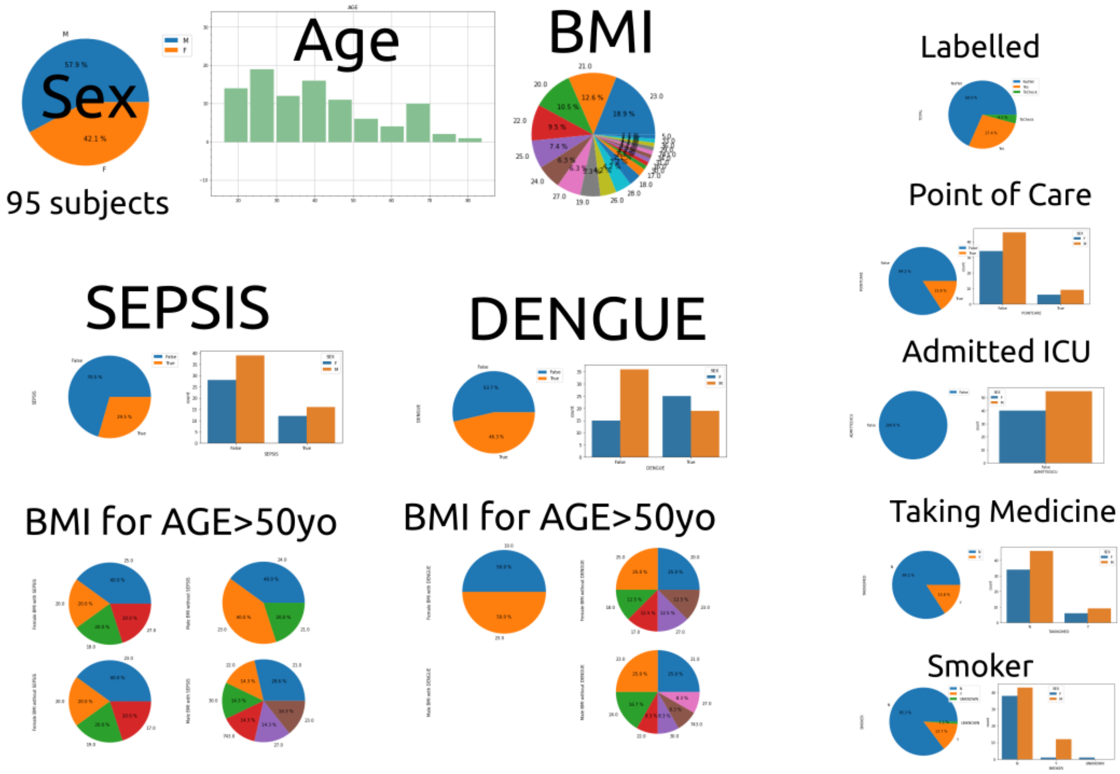

Figure 2: Patient demographics (**a**) description... (**b**) description... Figure is adapted from the works of

10 Potential future work

2D velocity vector fields of flow blow can help to detect abnormal flow patterns as done in fetal and neonatal echochardiograhy [Meyers et al., 2021]. Use LV A4C echos that can create synthetic Ulstraound images for GE Vivid E9, Hitachi Prosound U7, Philips iE 33 Vision, Siemens SC2000, and Toshiba Artida ultrasound systems [Brindise et al., 2020].

References

[Alessandrini et al., 2018] Alessandrini, M., Chakraborty, B., Heyde, B., Bernard, O., De Craene, M., Sermesant, M., and D'Hooge, J. (2018). Realistic vendor-specific synthetic ultrasound data for quality assurance of 2-d speckle tracking echocardiography: Simulation pipeline and open access database. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 65(3):411–422.

[Alessandrini et al., 2015] Alessandrini, M., Heyde, B., Giffard-Roisin, S., Delingette, H., Sermesant, M., Allain, P., Bernard, O., De Craene, M., and D'hooge, J. (2015). Generation of ultra-realistic synthetic echocardiographic sequences to facilitate standardization of deformation imaging. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 756–759.

[Asch et al., 2021] Asch, F. M., Mor-Avi, V., Rubenson, D., Goldstein, S., Saric, M., Mikati, I., Surette, S., Chaudhry, A., Poilvert, N., Hong, H., Horowitz, R., Park, D., Diaz-Gomez, J. L., Boesch, B., Nikravan, S., Liu, R. B., Philips, C., Thomas, J. D., Martin, R. P., and Lang, R. M. (2021). Deep learning-based automated echocardiographic quantification of left ventricular ejection fraction: A point-of-care solution. *Circulation: Cardiovascular Imaging*, 14(6):e012293.

[Asch et al., 2019] Asch, F. M., Poilvert, N., Abraham, T., Jankowski, M., Cleve, J., Adams, M., Romano, N., Hong, H., Mor-Avi, V., Martin, R. P., and Lang, R. M. (2019). Automated echocardiographic quantification of left ventricular ejection fraction without volume measurements using a machine learning algorithm mimicking a human expert. *Circulation: Cardiovascular Imaging*, 12(9):e009303.

[Aviles et al., 2017] Aviles, A. I., Widlak, T., Casals, A., Nillesen, M. M., and Ammari, H. (2017). Robust cardiac motion estimation using ultrafast ultrasound data: a low-rank topology-preserving approach. *Physics in Medicine and Biology*, 62(12):4831–4851.

[Bar et al., 2020] Bar, O., Neimark, D., Zohar, M., Hager, G. D., Girshick, R., Fried, G. M., Wolf, T., and Asselmann, D. (2020). Impact of data on generalization of ai for surgical intelligence applications. *Scientific Reports*, 10(1):22208.

[Baumgartner et al., 2017] Baumgartner, C. F., Kamnitsas, K., Matthew, J., Fletcher, T. P., Smith, S., Koch, L. M., Kainz, B., and Rueckert, D. (2017). Sononet: Real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE Transactions on Medical Imaging*, 36(11):2204–2215.

[Behnami et al., 2020] Behnami, D., Luong, C., Vaseli, H., Girgis, H., Abdi, A., Hawley, D., Gin, K., Rohling, R., Abolmaesumi, P., and Tsang, T. (2020). Automatic cine-based detection of patients at high risk of heart failure with reduced ejection fraction in echocardiograms. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 8(5):502–508.

[Bernard et al., 2016] Bernard, O., Bosch, J. G., Heyde, B., Alessandrini, M., Barbosa, D., Camarasu-Pop, S., Cervenansky, F., Valette, S., Mirea, O., Bernier, M., Jodoin, P.-M., Domingos, J. S., Stebbing, R. V., Keraudren, K., Oktay, O., Caballero, J., Shi, W., Rueckert, D., Milletari, F., Ahmadi, S.-A., Smistad, E., Lindseth, F., van Stralen, M., Wang, C., Smedby, Ö., Donal, E., Monaghan, M., Papachristidis, A., Geleijnse, M. L., Galli, E., and D'hooge, J. (2016). Standardized evaluation system for left ventricular segmentation algorithms in 3d echocardiography. *IEEE Transactions on Medical Imaging*, 35(4):967–977.

[Boice et al., 2022] Boice, E. N., Hernandez-Torres, S. I., and Snider, E. J. (2022). Comparison of ultrasound image classifier deep learning algorithms for shrapnel detection. *Journal of Imaging*, 8(5).

[Brindise et al., 2020] Brindise, M. C., Meyers, B. A., Kutty, S., and Vlachos, P. P. (2020). Unsupervised segmentation of b-mode echocardiograms.

[Campbell et al., 2018] Campbell, S. J., Bechara, R., and Islam, S. (2018). Point-of-care ultrasound in the intensive care unit. *Clinics in Chest Medicine*, 39(1):79–97. Interventional Pulmonology: An Update.

[Chartsias et al., 2021] Chartsias, A., Gao, S., Mumith, A., Oliveira, J., Bhatia, K., Kainz, B., and Beqiri, A. (2021). Contrastive learning for view classification of echocardiograms. In Noble, J. A., Aylward, S., Grimwood, A., Min, Z., Lee, S.-L., and Hu, Y., editors, *Simplifying Medical Ultrasound*, pages 149–158, Cham. Springer International Publishing.

[Cheema et al., 2021] Cheema, B. S., Walter, J., Narang, A., and Thomas, J. D. (2021). Artificial intelligence–enabled pocus in the covid-19 icu: A new spin on cardiac ultrasound. *JACC: Case Reports*, 3(2):258–263.

[Dong et al., 2020] Dong, S., Luo, G., Tam, C., Wang, W., Wang, K., Cao, S., Chen, B., Zhang, H., and Li, S. (2020). Deep atlas network for efficient 3d left ventricle segmentation on echocardiography. *Medical Image Analysis*, 61:101638.

[Dong et al., 2018] Dong, S., Luo, G., Wang, K., Cao, S., Mercado, A., Shmuilovich, O., Zhang, H., and Li, S. (2018). Voxelatlasgan: 3d left ventricle segmentation on echocardiography with atlas guided generation and voxel-to-voxel discrimination. In Frangi, A. F., Schnabel, J. A., Davatzikos, C., Alberola-López, C., and Fichtinger, G., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, pages 622–629, Cham. Springer International Publishing.

[Dukler et al., 2018] Dukler, Y., Ge, Y., Qian, Y., Yamamoto, S., Yuan, B., Zhao, L., Bertozzi, A. L., Hunter, B., Llerena, R., and Yen, J. T. (2018). Automatic valve segmentation in cardiac ultrasound time series data. In Angelini, E. D. and Landman, B. A., editors, *Medical Imaging 2018: Image Processing*, volume 10574, pages 493 – 504. International Society for Optics and Photonics, SPIE.

[Evain et al., 2022] Evain, E., Sun, Y., Faraz, K., Garcia, D., Saloux, E., Gerber, B. L., De Craene, M., and Bernard, O. (2022). Motion estimation by deep learning in 2d echocardiography: synthetic dataset and validation. *IEEE Transactions on Medical Imaging*, pages 1–1.

[Feigenbaum, 1996] Feigenbaum, H. (1996). Evolution of echocardiography. *Circulation*, 93(7):1321–1327.

[Field et al., 2011] Field, L. C., Guldan, G. J., and Finley, A. C. (2011). Echocardiography in the intensive care unit. *Seminars in Cardiothoracic and Vascular Anesthesia*, 15(1-2):25–39. PMID: 21719547.

[Francesconi et al., 2021] Francesconi, M., Martina, M. R., Armenia, S., Buzzelli, A., Di Franco, G., Gemignani, V., Bianchini, E., and Bruno, R. M. (2021). Technical validation and usability of a portable ultrasound-based system for carotid assessment of vascular ageing: A pilot study. *Heart, Lung and Circulation*, 30(11):1734–1743.

[Ghorbani et al., 2020] Ghorbani, A., Ouyang, D., Abid, A., He, B., Chen, J. H., Harrington, R. A., Liang, D. H., Ashley, E. A., and Zou, J. Y. (2020). Deep learning interpretation of echocardiograms. *npj Digital Medicine*, 3(1):10.

[Hanson III and Marshall, 2001] Hanson III, C. W. and Marshall, B. E. (2001). Artificial intelligence applications in the intensive care unit. *Critical Care Medicine*, 29(2).

[Hao et al., 2021] Hao, N., Yen, L., Davies-Foote, R., Trung, T., Duoc, N., Trang, V., Nhat, P., Duc, D., Anh, N., Lieu, P., Thuy, T., Thuy, D., Phong, N., Truong, N., Thanh, P., Tam, D., Puthucheary, Z., and Thwaites, C. (2021). The management of tetanus in adults in an intensive care unit in southern vietnam [version 2; peer review: 3 approved]. *Wellcome Open Research*, 6(107).

[Hong et al., 2021] Hong, H., Surette, S., Chaudhry, A. K., Parajuli, N., Cadieu, C., Martin, R., and Thomas, J. (2021). Ai-guided echocardiography system matches the image quality assessment ability of cardiac sonographers. *Journal of the American College of Cardiology*, 77(18 Supplement 1):3240–3240.

[Hong et al., 2022] Hong, N., Liu, C., Gao, J., Han, L., Chang, F., Gong, M., and Su, L. (2022). State of the art of machine learning–enabled clinical decision support in intensive care units: Literature review. *JMIR Med Inform*, 10(3):e28781.

[Jafari et al., 2021] Jafari, M. H., Woudenberg, N. V., Luong, C., Abolmaesumi, P., and Tsang, T. (2021). Deep bayesian image segmentation for a more robust ejection fraction estimation. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 1264–1268.

[Khamis et al., 2017] Khamis, H., Zurakhov, G., Azar, V., Raz, A., Friedman, Z., and Adam, D. (2017). Automatic apical view classification of echocardiograms using a discriminative learning dictionary. *Medical Image Analysis*, 36:15–21.

[Kusunose, 2021] Kusunose, K. (2021). Steps to use artificial intelligence in echocardiography. *Journal of Echocardiography*, 19(1):21–27.

[Labs et al., 2021] Labs, R. B., Zolgharni, M., and Loo, J. P. (2021). Echocardiographic image quality assessment using deep neural networks. In Papież, B. W., Yaqub, M., Jiao, J., Namburete, A. I. L., and Noble, J. A., editors, *Medical Image Understanding and Analysis*, pages 488–502, Cham. Springer International Publishing.

[Laumer et al., 2020] Laumer, F., Fringeli, G., Dubatovka, A., Manduchi, L., and Buhmann, J. M. (2020). Deepheartbeat: Latent trajectory learning of cardiac cycles using cardiac ultrasounds. In Alsentzer, E., McDermott, M. B. A., Falck, F., Sarkar, S. K., Roy, S., and Hyland, S. L., editors, *Proceedings of the Machine Learning for Health NeurIPS Workshop*, volume 136 of *Proceedings of Machine Learning Research*, pages 194–212. PMLR.

[Leclerc et al., 2019] Leclerc, S., Smistad, E., Pedrosa, J., Østvik, A., Cervenansky, F., Espinosa, F., Espeland, T., Berg, E. A. R., Jodoin, P.-M., Grenier, T., Lartizien, C., D'hooge, J., Lovstakken, L., and Bernard, O. (2019). Deep learning for segmentation using an open large-scale dataset in 2d echocardiography. *IEEE Transactions on Medical Imaging*, 38(9):2198–2210.

[Liu et al., 2021] Liu, F., Wang, K., Liu, D., Yang, X., and Tian, J. (2021). Deep pyramid local attention neural network for cardiac structure segmentation in two-dimensional echocardiography. *Medical Image Analysis*, 67:101873.

[Lu et al., 2020] Lu, P., Qiu, H., Qin, C., Bai, W., Rueckert, D., and Noble, J. A. (2020). Going deeper into cardiac motion analysis to model fine spatio-temporal features. In

Papież, B. W., Namburete, A. I. L., Yaqub, M., and Noble, J. A., editors, *Medical Image Understanding and Analysis*, pages 294–306, Cham. Springer International Publishing.

[Meyers et al., 2021] Meyers, B. A., Brindise, M. C., Payne, R. M., and Vlachos, P. P. (2021). An integrated and automated tool for quantification of biomechanics in fetal and neonatal echocardiography. *medRxiv*.

[Mohamed et al., 2021] Mohamed, S., Muhtaseb, R., and Mohammad, Y. (2021). Is contrastive learning suitable for left ventricular segmentation in echocardiographic images?

[Narang et al., 2021] Narang, A., Bae, R., Hong, H., Thomas, Y., Surette, S., Cadieu, C., Chaudhry, A., Martin, R. P., McCarthy, P. M., Rubenson, D. S., Goldstein, S., Little, S. H., Lang, R. M., Weissman, N. J., and Thomas, J. D. (2021). Utility of a Deep-Learning Algorithm to Guide Novices to Acquire Echocardiograms for Limited Diagnostic Use. *JAMA Cardiology*, 6(6):624–632.

[Ono et al., 2022] Ono, S., Komatsu, M., Sakai, A., Arima, H., Ochida, M., Aoyama, R., Yasutomi, S., Asada, K., Kaneko, S., Sasano, T., and Hamamoto, R. (2022). Automated endocardial border detection and left ventricular functional assessment in echocardiography using deep learning. *Biomedicines*, 10(5).

[Ouyang et al., 2019] Ouyang, D., He, B., Ghorbani, A., Matt, L. P., A., A. E., H., L. D., and Y., Z. J. (2019). Echonet-dynamic: a large new cardiac motion video data resource for medical machine learing.

[Ouyang et al., 2020] Ouyang, D., He, B., Ghorbani, A., Yuan, N., Ebinger, J., Langlotz, C. P., Heidenreich, P. A., Harrington, R. A., Liang, D. H., Ashley, E. A., and Zou, J. Y. (2020). Video-based ai for beat-to-beat assessment of cardiac function. *Nature*, 580(7802):252–256.

[Reynaud et al., 2021] Reynaud, H., Vlontzos, A., Hou, B., Beqiri, A., Leeson, P., and Kainz, B. (2021). Ultrasound video transformers for cardiac ejection fraction estimation. In de Bruijne, M., Cattin, P. C., Cotin, S., Padoy, N., Speidel, S., Zheng, Y., and Essert, C., editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, pages 495–505, Cham. Springer International Publishing.

[Rubin et al., 2021] Rubin, J., Erkamp, R., Naidu, R. S., Thodiyil, A. O., and Chen, A. (2021). Attention distillation for detection transformers: Application to real-time video object detection in ultrasound. In Roy, S., Pfohl, S., Rocheteau, E., Tadesse, G. A., Oala, L., Falck, F., Zhou, Y., Shen, L., Zamzmi, G., Mugambi, P., Zirikly, A., McDermott, M. B. A., and Alsentzer, E., editors, *Proceedings of Machine Learning for Health*, volume 158 of *Proceedings of Machine Learning Research*, pages 26–37. PMLR.

[Salte et al., 2021] Salte, I. M., Østvik, A., Smistad, E., Melichova, D., Nguyen, T. M., Karlsen, S., Brunvand, H., Haugaa, K. H., Edvardsen, T., Lovstakken, L., and Grenne, B. (2021). Artificial intelligence for automatic measurement of left ventricular strain in echocardiography. *JACC: Cardiovascular Imaging*.

[Singh and Goyal, 2007] Singh, S. and Goyal, A. (2007). The origin of echocardiography: a tribute to inge edler. *Texas Heart Institute journal*, 34(4):431–438.

[Smistad et al., 2021a] Smistad, E., Salte, I. M., Dalen, H., and Lovstakken, L. (2021a). Real-time temporal coherent left ventricle segmentation using convolutional lstms. In *2021 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4.

[Smistad et al., 2021b] Smistad, E., Steinsland, E. N., and Løvstakken, L. (2021b). Real-time 3d left ventricle segmentation and ejection fraction using deep learning. In *2021 IEEE International Ultrasonics Symposium (IUS)*, pages 1–3.

[Smistad et al., 2021c] Smistad, E., Østvik, A., and Løvstakken, L. (2021c). Annotation web - an open-source web-based annotation tool for ultrasound images. In *2021 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4.

[Snider et al., 2022] Snider, E. J., Hernandez-Torres, S. I., and Boice, E. N. (2022). An image classification deep-learning algorithm for shrapnel detection from ultrasound images. *Scientific Reports*, 12(1):8427.

[Toussaint et al., 2018] Toussaint, N., Khanal, B., Sinclair, M., Gomez, A., Skelton, E., Matthew, J., and Schnabel, J. A. (2018). Weakly supervised localisation for fetal ultrasound images. In Stoyanov, D., Taylor, Z., Carneiro, G., Syeda-Mahmood, T., Martel, A., Maier-Hein, L., Tavares, J. M. R., Bradley, A., Papa, J. P., Belagiannis, V., Nascimento, J. C., Lu, Z., Conjeti, S., Moradi, M., Greenspan, H., and Madabhushi, A., editors, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 192–200, Cham. Springer International Publishing.

[Tromp et al., 2022] Tromp, J., Seekings, P. J., Hung, C.-L., Iversen, M. B., Frost, M. J., Ouwerkerk, W., Jiang, Z., Eisenhaber, F., Goh, R. S. M., Zhao, H., Huang, W., Ling, L.-H., Sim, D., Cozzone, P., Richards, A. M., Lee, H. K., Solomon, S. D., Lam, C. S. P., and Ezekowitz, J. A. (2022). Automated interpretation of systolic and diastolic function on the echocardiogram: a multicohort study. *The Lancet Digital Health*, 4(1):e46–e54.

[Ulloa Cerna et al., 2021] Ulloa Cerna, A. E., Jing, L., Good, C. W., vanMaanen, D. P., Raghunath, S., Suever, J. D., Nevius, C. D., Wehner, G. J., Hartzel, D. N., Leader, J. B., Alsaid, A., Patel, A. A., Kirchner, H. L., Pfeifer, J. M., Carry, B. J., Pattichis, M. S., Haggerty, C. M., and Fornwalt, B. K. (2021). Deep-learning-assisted analysis of echocardiographic videos improves predictions of all-cause mortality. *Nature Biomedical Engineering*, 5(6):546–554.

[Van Woudenberg et al., 2018] Van Woudenberg, N., Liao, Z., Abdi, A. H., Girgis, H., Luong, C., Vaseli, H., Behnami, D., Zhang, H., Gin, K., Rohling, R., Tsang, T., and Abolmaesumi, P. (2018). Quantitative echocardiography: Real-time quality estimation and view classification implemented on a mobile android device. In Stoyanov, D., Taylor, Z., Aylward, S., Tavares, J. M. R., Xiao, Y., Simpson, A., Martel, A., Maier-Hein, L., Li, S., Rivaz, H., Reinertsen, I., Chabanas, M., and Farahani, K., editors, *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*, pages 74–81, Cham. Springer International Publishing.

[Vieillard-Baron et al., 2008] Vieillard-Baron, A., Slama, M., Cholley, B., Janvier, G., and Vignon, P. (2008). Echocardiography in the intensive care unit: from evolution to revolution? *Intensive Care Medicine*, 34(2):243–249.

[Wu et al., 2022] Wu, H., Liu, J., Xiao, F., Wen, Z., Cheng, L., and Qin, J. (2022). Semi-supervised segmentation of echocardiography videos via noise-resilient spatiotemporal semantic calibration and fusion. *Medical Image Analysis*, 78:102397.

[Xie et al., 2020] Xie, J., Song, X., Zhang, W., Dong, Q., Wang, Y., Li, F., and Wan, C. (2020). A novel approach with dual-sampling convolutional neural network for ultrasound image classification of breast tumors. *Physics in Medicine and Biology*, 65(24):245001.

[Young et al., 2021] Young, C., Ahn, , 20983, , Ahn, C. Y., Sangwoon, Yun, , 20984, , and Yun, S. (2021). Nonconvex constrained minimisation for 3d left ventricular shape recovery using 2d echocardiography data. *East Asian Journal on Applied Mathematics*, 12(1):111–124.

[Yuan et al., 2017] Yuan, B., Chitturi, S. R., Iyer, G., Li, N., Xu, X., Zhan, R., Llerena, R., Yen, J. T., and Bertozzi, A. L. (2017). Machine learning for cardiac ultrasound time series data. In Krol, A. and Gimi, B., editors, *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 10137, pages 617 – 624. International Society for Optics and Photonics, SPIE.

[Zhang et al., 2018] Zhang, J., Gajjala, S., Agrawal, P., Tison, G. H., Hallock, L. A., Beussink-Nelson, L., Lassen, M. H., Fan, E., Aras, M. A., Jordan, C., Fleischmann, K. E., Melisko, M., Qasim, A., Shah, S. J., Bajcsy, R., and Deo, R. C. (2018). Fully automated echocardiogram interpretation in clinical practice. *Circulation*, 138(16):1623–1635.

[Zhang et al., 2022] Zhang, Z., Zhu, Y., Liu, M., Zhang, Z., Zhao, Y., Yang, X., Xie, M., and Zhang, L. (2022). Artificial intelligence-enhanced echocardiography for systolic function assessment. *Journal of Clinical Medicine*, 11(10).

[Østvik et al., 2021] Østvik, A., Salte, I. M., Smistad, E., Nguyen, T. M., Melichova, D., Brunvand, H., Haugaa, K., Edvardsen, T., Grenne, B., and Lovstakken, L. (2021). Myocardial function imaging in echocardiography using deep learning. *IEEE Transactions on Medical Imaging*, 40(5):1340–1351.