

Artificial Intelligence assisted image extraction in neonatal echocardiography in Sub-Saharan Africa: A Protocol

Aminkeng Zawuo Leke, Lionel Landry Sop Deffo, Yunkavi Sabastian Wirsiy, Thomas Aldersley, Thomas Day, Andrew P King, Patrick McAllister, Michel N Maboh, John Lawrenson, Cabral Tantchou, Bernhard Kainz, Frank Casey, Raymond Bond, Finlay Dewar, Ngoe Kelson Tchinda, Obale Armstrong, Frunwi Ndeh Mugri, Liesl Zühlke, Helen Dolk

Submitted to: JMIR Research Protocols
on: April 14, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 28

 Figures 29

 Figure 1..... 30

 Figure 2..... 31

 Figure 3..... 32

 Figure 4..... 33

 Figure 5..... 34

 Figure 6..... 35

 Figure 7..... 36

Artificial Intelligence assisted image extraction in neonatal echocardiography in Sub-Saharan Africa: A Protocol

Aminkeng Zawuo Leke^{1*}; Lionel Landry Sop Deffo^{1*}; Yunkavi Sabastian Wirsy^{2*}; Thomas Aldersley^{3*}; Thomas Day^{4*}; Andrew P King^{4*}; Patrick McAllister^{5*}; Michel N Maboh^{1*}; John Lawrenson^{3*}; Cabral Tantchou^{6*}; Bernhard Kainz^{7*}; Frank Casey^{8*}; Raymond Bond^{5*}; Finlay Dewar^{5*}; Ngoe Kelson Tchinda^{1*}; Obale Armstrong^{2*}; Frunwi Ndeh Mugri^{1*}; Liesl Zühlke^{3*}; Helen Dolk^{8*}

¹Great Soppo Buea Digital Technology and Innovation Hub, Health Research Foundation Buea Buea CM

²Digital Technology and Innovation Hub, Health Research Foundation Buea Buea CM

³South African Medical Research Council Department of Paediatrics and Child Health, Division of Paediatric Cardiology, University of Cape Town Cape Town ZA

⁴School of Biomedical Engineering and Imaging Sciences, King's College London London GB

⁵School of Engineering, Belfast Campus, Ulster University Belfast GB

⁶Digital Technology and Innovation Hub, Health Research Foundation Buea Shisong cardiac centre / St Padre Pio hospital Douala CM

⁷Department of Computing, South Kensington Campus, Imperial College London London GB

⁸School of Medicine, Ulster University Londonderry GB

*these authors contributed equally

Corresponding Author:

Aminkeng Zawuo Leke

Great Soppo Buea

Digital Technology and Innovation Hub, Health Research Foundation Buea

Compain Street, Great Soppo

Buea

CM

Abstract

Background: Sub-Saharan Africa (SSA) has the highest global burden of under-five child mortality, with congenital heart disease (CHD) being a major contributor. Despite advancements in high-income countries, CHD-related mortality remains unchanged in SSA due to limited diagnostic capacity and centralized healthcare. While pulse oximetry supports early detection, confirmation of diagnosis often relies on echocardiography, a procedure hindered by a shortage of specialized personnel. Artificial intelligence (AI) offers a promising solution to address this diagnostic gap.

Objective: This study aims to develop an AI-assisted echocardiography system that will enable non-expert operators such as nurses midwives and medical doctors to perform basic cardiac ultrasound sweeps on neonates suspected of CHD and extract accurate cardiac images that can be sent to a remote paediatric cardiologist for interpretation

Methods: The study will follow a two-phase approach to develop a deep learning model for real-time cardiac view detection in neonatal echocardiography, using data from St Padre Pio Hospital in Cameroon and the Red Cross War Memorial Children's Hospital in South Africa ensuring demographic diversity. Phase one will pretrain the model on retrospective data from ~500 neonates (0–28 days old). Phase two will fine-tune it using prospective data from 1,000 neonates, which includes background elements absent in the retrospective set, enabling adaptation to local clinical environments. The datasets will include short and continuous echocardiographic video clips covering ten standard cardiac views, as defined by the American Society of Echocardiography. The model architecture will leverage convolutional neural networks (CNNs) and convolutional Long Short-Term Memory (convLSTM) layers, inspired by the interleaved visual memory framework, which combines fast and slow feature extractors through a shared temporal memory mechanism. Video preprocessing, annotation with predefined cardiac view codes using Labelbox, and training with TensorFlow and PyTorch will be conducted. Reinforcement learning will guide the dynamic use of feature extractors during training. Iterative refinement, supported by clinical input, will ensure the model effectively distinguishes correct from incorrect views in real-time, enhancing usability in resource-limited settings.

Results: Retrospective data collection for the project began in September 2024, and since then, data from 308 babies have been collected and labelled. In parallel, the initial model framework has been developed and training initiated using a small portion of

the labelled data. The project is currently at the intensive execution phase with all objectives running in parallel and final results expected within 14 months.

Conclusions: The AI-assisted echocardiography model that will be developed in this project holds promise for improving early CHD diagnosis and care in SSA and other low resource settings.

(JMIR Preprints 14/04/2025:75270)

DOI: <https://doi.org/10.2196/preprints.75270>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in a JMIR journal, my accepted manuscript PDF will be visible to all users.

No. Please do not make my accepted manuscript PDF available to anyone.

Original Manuscript

Artificial Intelligence assisted image extraction in neonatal echocardiography in Sub-Saharan Africa: A Protocol

Abstract

Background: Sub-Saharan Africa (SSA) has the highest global burden of under-five child mortality, with congenital heart disease (CHD) being a major contributor. Despite advancements in high-income countries, CHD-related mortality remains unchanged in SSA due to limited diagnostic capacity and centralized healthcare. While pulse oximetry supports early detection, confirmation of diagnosis often relies on echocardiography, a procedure hindered by a shortage of specialized personnel. Artificial intelligence (AI) offers a promising solution to address this diagnostic gap.

Objective: This study aims to develop an AI-assisted echocardiography system that will enable non-expert operators such as nurses midwives and medical doctors to perform basic cardiac ultrasound sweeps on neonates suspected of CHD and extract accurate cardiac images that can be sent to a remote paediatric cardiologist for interpretation.

Methods: The study will follow a two-phase approach to develop a deep learning model for real-time cardiac view detection in neonatal echocardiography, using data from St Padre Pio Hospital in Cameroon and the Red Cross War Memorial Children's Hospital in South Africa ensuring demographic diversity. Phase one will pretrain the model on retrospective data from ~500 neonates (0–28 days old). Phase two will fine-tune it using prospective data from 1,000 neonates, which includes background elements absent in the retrospective set, enabling adaptation to local clinical environments. The datasets will include short and continuous echocardiographic video clips covering ten standard cardiac views, as defined by the American Society of Echocardiography. The model architecture will leverage convolutional neural networks (CNNs) and convolutional Long Short-Term Memory (convLSTM) layers, inspired by the interleaved visual memory framework, which combines fast and slow feature extractors through a shared temporal memory mechanism. Video preprocessing, annotation with predefined cardiac view codes using Labelbox, and training with TensorFlow and PyTorch will be conducted. Reinforcement learning will guide the dynamic use of feature extractors during training. Iterative refinement, supported by clinical input, will ensure the model effectively distinguishes correct from incorrect views in real-time, enhancing usability in resource-limited settings.

Results: Retrospective data collection for the project began in September 2024, and since then, data from 308 babies have been collected and labelled. In parallel, the initial model framework has been developed and training initiated using a small portion of the labelled data. The project is currently at the intensive execution phase with all objectives running in parallel and final results expected within 14 months.

Conclusions: The AI-assisted echocardiography model that will be developed in this project holds promise for improving early CHD diagnosis and care in SSA and other low resource settings.

Keywords: AI-assisted echocardiography; Congenital heart defect (CHD) screening; Neonatal cardiac imaging; Sub-Saharan Africa healthcare; Telemedicine and AI integration

Introduction

Sub-Saharan Africa (SSA) is the region of the world with the highest burden of under-five child

mortality – 56.7% of the 4.9 million global under-five deaths were in SSA in 2022 [1], compared to only one quarter of births. As progress has been made tackling infectious diseases in children, the relative burden of other causes of death, including congenital anomalies (CA), is increasing, such that CA are now the sixth highest cause of under-five death in SSA [1]. The largest contributor to CA, both in terms of prevalence at birth and mortality, is congenital heart disease (CHD). While child mortality associated with CHD in most high-income countries (HICs) has been declining, often dramatically, there has been no improvement in most of SSA [2]. This is impeding progress towards SDG3.2, a reduction in child mortality, as well as SDG3.8, equity of access to quality essential health-care services.

The challenges of CHD care in SSA are multifactorial. A recent analysis of African paediatric and congenital heart disease services showed that only 18 (40%) African countries were able to provide a full CHD service including cardiac surgery and interventional cardiac catheterization [3]. Furthermore, the capacity of these services is low, with poor infrastructure and severe shortages of specialized personnel. Indeed, there are only 0.17 (IQR: 0.02 – 0.35) paediatric cardiologists per million population, which is far below the international recommended ratio of 2 per million population [3].

This lack of capacity results in late and missed diagnoses, leading to delayed intervention, negatively affecting outcomes. [4,5] This is compounded by the centralization of paediatric cardiac facilities to major urban areas, often hundreds of kilometers away from most birthing centres [4]. To address these challenges, interventions that enhance the capacity for early CHD detection and decentralize care are critical.

Postnatal screening of CHD involves a combination of clinical examination, pulse oximetry screening and echocardiography. Pulse oximetry screening is a simple, non-invasive test that measures the oxygen saturation in a newborn's blood [6] Low oxygen levels or a significant difference between readings in the hand and foot may indicate critical CHD, prompting further investigation with echocardiography. Despite some recent concerns about the suboptimal performance of the device in dark-skinned infants [7,8] pulse oximetry screening is highly recommended and is currently being adopted by many LMICs, including in SSA.[6,9] However, pulse oximetry screening results are nonspecific to CHD, as low oxygen saturation could also be an indication of other conditions such pneumonia - the leading cause of neonatal mortality in SSA. Therefore, all suspected CHD cases from pulse oximetry must be referred to an expert paediatric cardiologist for confirmation through echocardiography. Consequently, the increased uptake of pulse oximetry in SSA is increasing the number of suspected CHD referrals and the workload for the already strained workforce of echocardiographers and paediatric cardiologists. This is further compounded by the cost involved and the arduous long journey that must be undertaken by the particularly fragile neonates, some of whom may eventually turn out to be false positive cases.

The introduction of low-cost compact mobile ultrasound devices has dramatically increased the availability of ultrasound services in low- and middle-income countries (LMICs) and

gone some way to address the centralization of diagnostic services [10]. However, the lack of diagnostic expertise still limits progress in using these devices to perform highly specialized procedures such as echocardiography. Obtaining optimal echocardiographic images presents significant challenges [10] The echocardiogram probe must be properly positioned, and parameters such as gain, contrast, resolution, depth, and magnification must be individually set for each specific view. [11] When combined with patient factors like uncooperative pediatric behavior and complicated CHD anatomy, these factors lead to long scan times, increased workloads, and the need for highly trained sonographers or echocardiographers.

Training programs have been shown to enhance both image capture and the identification of anomalies [11] However, these programs are often labor-intensive and time-consuming, requiring repetition due to staff turnover [11] Therefore, a complementary approach is necessary to maintain and build upon the improvements achieved through training.

In recent years there have been a growing number of projects using artificial intelligence (AI) models to facilitate CHD diagnosis. [11,12,13,14,15] However, these projects primarily originate from HICs, potentially limiting their applicability in LMICs due to domain shifts related to equipment, operator expertise, ethnicity, and disease profiles. Moreover, existing research disproportionately focuses on adult echocardiography, with limited attention to prenatal CHD and a neglect of neonatal applications. Critically, current AI efforts center on image interpretation, rather than assisting non-experts with real-time image acquisition.

This paper outlines an ongoing project leveraging AI to enable non-expert echocardiographers—such as sonographers, nurses, midwives, and doctors with minimal specialized training—to perform basic cardiac ultrasound "sweeps" on neonates (aged 0–28 days). The AI system will be designed to automatically extract predefined cardiac views, which can then be transmitted to remote paediatric cardiologists for interpretation. The AI model will be limited to image extraction and will not extend to diagnosing or identifying potential CHD. This approach aims to streamline the initial screening process for CHD. The AI-assisted echocardiography will be selectively used for babies who have shown positive results in pulse oximetry screening or presented with other clinical indications suggestive of potential cardiac issues, helping to identify cases that may require further evaluation by specialised pediatric cardiologists (Figure 1)

Figure 1: AI assisted echocardiography CHD screening process

Methods

Study population and cardiac views

This study focuses on extracting accurate cardiac views in neonates aged 0-28 days suspected of CHD. To guide the selection of relevant cardiac views, two critical questions were considered: (1) whether non-experts could reliably maneuver the probe to enable the

AI system to capture the desired views, and (2) which cardiac views would provide the most valuable insights for cardiologists in determining the need for further investigation. Based on these considerations and in alignment with the 2024 American Society of Echocardiography (ASE) guidelines [16], our panel of expert cardiologists identified 10 essential cardiac views (Table 1).

Table 1 Selected cardiac views of interest and label codes

SN	Cardiac View	Label Code
1	Subcostal long axis	SCLA
2	Apical 4 chamber	A4C
3	Apical 5 chamber	A5C
4	Parasternal long axis	PLAX
5	Parasternal short axis – Aortic Valve	PSAX_AV
6	Parasternal short axis – Mitral Valve	PSAX_MV
7	Parasternal short axis – Papillary Muscle	PSAX_PM
8	Suprasternal Long Axis	SSLAX
9	Other diagnostic	OTHER_DIAG
10	Background	BKG

Data requirements

To train AI models that guide non-expert operators in acquiring specific cardiac views, a dataset of real patient ultrasound videos is essential. This dataset must include not only the desired (positive) views but also a substantial amount of background (negative) data, representing non-essential or incorrect views. This is crucial for the model to effectively learn to distinguish target views from the more common, non-ideal images that non-expert operators are likely to produce in real-world settings.

Existing data saved by expert cardiologists during routine clinical scans provide a readily available source with limited ethical concerns. However, cardiologists typically save only the most relevant cardiac views during a scan, meaning that existing hospital datasets contain only these specific views without the necessary background elements. Additionally, a

significant domain shift exists between existing datasets—primarily collected in high-income countries—and newly acquired data from low- and middle-income countries (LMICs), where variations in clinical protocols, imaging equipment, and patient demographics may impact model generalizability. Furthermore, there is a lack of publicly available neonate echocardiography datasets, making data acquisition a critical challenge. To obtain data that contains both the target views and background elements, new data must be collected prospectively. This approach also enables modifications to facilitate frame extraction and labeling, as subsequently described in the Labelling section. However, collecting prospective data is time-consuming and resource-intensive. Ethical considerations must also be addressed, including obtaining informed consent from parents and ensuring that data collection does not disrupt routine clinical workflows. A sensible approach to mitigate these limitations while maximizing the advantages of both dataset types is to develop the model in two phases: initially leveraging existing hospital datasets for pretraining and subsequently fine-tuning the model with prospectively collected data to ensure adaptability to LMIC settings.

In the first phase, existing data will be used to build an initial model framework, while simultaneously collecting the prospective data. Once the new prospective data becomes available, it will be used in the second phase to upgrade and refine the initial model that was developed using the existing data (see section on model development). This hybrid approach allows the project to make the most of the readily available existing data to kickstart the model development, while also ensuring the final model is trained on a comprehensive dataset that includes the target cardiac views as well as the necessary background elements. Both types of datasets will include all scanned cases, regardless of whether the baby's heart was abnormal.

The proposed AI model is meant for use in neonates aged 0-28 days. However, to expedite the data collection process and increase the sample size, the project will include data from babies aged 0-2 years, based on the observation that the heart structure of older infants is similar to neonates [17].

In the first phase of the model-building process, the project will aim to gather short anonymized echocardiographic video clips from 500 babies. The dataset will consist of 1 to 7 short video clips per baby, depending on the availability and quality of the scans. These clips are expected to include one or more of the 10 cardiac views utilized in the project. Each video clip is projected to be approximately 3-5 seconds long and will have an average file size of up to 4MB. Based on these parameters, the total number of clips is estimated to range from 500 to 3,500, resulting in an expected storage requirement of around 14GB for the raw video data. Additional space for data preprocessing, model training, and backups will increase the total storage needed for this phase to approximately 200GB, including 50GB allocated for backups.

In the second phase of the model development, the project will collect prospective echocardiographic video data from 1,000 babies. Each baby will have one continuous video clip representing a complete echocardiographic scan, which will be recorded at a higher frame rate of 60 frames per second, with an average duration of 20 minutes. Shorter video clips corresponding to individual cardiac views will subsequently be extracted from these long recordings for analysis. It is anticipated that each 20-minute video will have an average file size of 1GB, resulting in a total raw

data size of approximately 1TB. With data preprocessing, the extraction of shorter clips, model training, and storage redundancy, the total storage requirement for this phase is projected to reach around 15TB, including 5TB allocated for backups.

By focusing on video clips directly, the project will simplify data collection while preserving the variability and dynamic features necessary for robust model training. The shorter clips extracted in the second phase will ensure compatibility with the analysis framework established during the first phase.

Study sites and data collection

The project will collect data from two hospitals – St Padre Pio Hospital in Douala Cameroon and the Red Cross War Memorial Children's Hospital in Cape Town, South Africa. About 80% of the data will come from South Africa while Cameroon will contribute 20% of the data. These sites have been carefully selected to ensure representation of diverse patient demographics, which will be critical for the performance and external validity of the developed model.

Cameroon is a bilingual (French and English) country located in central Africa and often termed "Africa in Miniature". The country captures the vast diversity in SSA, encompassing multiple cultural, linguistic, and geographical variations. St Padre Pio Hospital in Douala is one of the few referral hospitals for paediatric cardiology in Cameroon and receives patients from across the country. The Children's Heart Disease Research Unit at the Red Cross War Memorial Children's Hospital in Cape Town serves a racially and ethnically diverse population, with considerable variability in socioeconomic factors and healthcare access. Together, data from Cameroon and South Africa will provide the necessary representativeness to ensure the model performs effectively across different clinical conditions and patient demographics in SSA.

Existing data (short video clips) will be extracted from the local saved sources at the participating hospitals. The prospective data will be collected during routine echocardiography scans performed by expert cardiologists or sonographers. Prior to the data collection, the experts will obtain informed consent from the parents or legal guardians of the participating infant. The informed consent form and information sheet detailing the study are provided in Appendix A. Parents will have the choice to refuse participation in the study without any impact on the care the child receives. The project team is committed to ensuring the data collection is conducted in an ethical manner that respects the rights and well-being of the participating families. It is crucial that the data collection process does not interfere with the normal clinical care of the infants. The expert is expected to conduct the scan as usual with two simple modifications:

- 1) prior to starting the scan, the expert shall connect a screen capture device to capture the entire echo scan from start to finish,
- 2) for the cardiac views of interest, in addition to the usual routine pause to save the frame, the expert is expected to also label the frame at a defined position (top right corner of the screen).

At the end of the scan, the captured video of the entire scan will be saved in the separate screen capture device. An SOP has been developed to guide all the cardiologist involved in the prospective data collection, minimizing operator variability (Appendix B). Our cardiologist experts believe that labeling the frames during the scan will not influence the typical scan time.

Data De-identification, Metadata Generation, and Data Transmission & Storage

The current project is part of the broader Data Science for Health Discovery and Innovation in Africa (DS-I Africa) Network. As one of the network's research projects, the project will leverage on the on-premises cloud storage and compute infrastructure provided by DS-I Africa through its eLwazi Open Data Science Platform (eLwazi ODSP). This infrastructure, specifically the Illifu system, forms the backbone of the data transmission and storage processes described in the Data Transmission and Storage section below, enabling secure and efficient data management.

All existing and prospective data will be transmitted to a central, secure on-premises cloud storage solution, where model building and data analysis will take place. Prior to transmission, the data will undergo several processes, including de-identification (encompassing both anonymisation and pseudonymisation), metadata generation, and secure transmission. These processes will ensure and guarantee data privacy, integrity, and security are maintained while adhering to the highest standards of data protection.

The thematic diagram (Figure 2) below illustrates the overall process flow, including data collection, de-identification, metadata generation, and secure transmission and storage. This diagram provides a visual overview of how data transitions through each stage of the workflow.

Figure 2 Thematic data flow diagram for CHD-AI project

1. Data De-identification

To ensure patient privacy and compliance with ethical guidelines, all data will undergo a de-identification process, removing personally identifiable information (PII) before use. This process will be implemented using a custom anonymization script developed by the project engineering team, which automates the removal of sensitive details while maintaining data integrity. Additionally, unique project identification codes will be assigned to facilitate internal traceability without exposing personal information. In cases where controlled data access is required for specific research purposes, a secure pseudonymization approach will be used, allowing authorized personnel to reference anonymized data without compromising privacy

2. Metadata Generation

After de-identification, metadata is generated to provide contextual information about the data, as shown in (Figure 2). This metadata will support future management and analysis.

Types of metadata: Includes patient's age, sex, and diagnosis that will support analysis and traceability.

Automated metadata extraction: Ensures consistency and minimizes human error by automating metadata generation with the aid of a script built by the engineering

team.

Use of metadata: Metadata will accompany anonymised and pseudonymised data as needed, maintaining transparency and traceability while adhering to privacy requirements.

3. Data Transmission and Storage

For this study, data will be securely transmitted to the ILLIFU on-premises cloud storage solution provided by the eLwazi Open Data Science Platform (eLwazi ODSP), a critical infrastructure component of the DS-I Africa Network. This stage, as illustrated in (Figure 2), ensures private and compliant storage accessible only to authorized personnel. The platform will also serve as the compute environment for model building, offering the necessary resources for data processing and analysis.

The decision to utilize the Illifu infrastructure stems from its robust integration into DS-I Africa's goals of fostering advanced data science capabilities across its projects. While Illifu forms the primary infrastructure, similar objectives can be achieved through other suitable cloud storage solutions, such as Google Cloud Platform (GCP), Amazon Web Services (AWS), or Microsoft Azure, provided they are configured to meet necessary privacy and compliance requirements.

Data labelling

1. Existing data labelling

The labeling of the existing echocardiographic data will be conducted by expert cardiologists using the Labelbox platform, integrated with the Google Cloud platform. This process will ensure consistent and high-quality annotations that are critical for model training.

a. Labeling Parameters

The labeling will involve the following parameters:

- i. Cardiac Views*
 - o Each video clip will be assigned one of the ten predefined cardiac views as outlined by the expert cardiologists in the project.
 - o For each view, a sub-parameter will indicate the Color Doppler presence, labeled as Color or No Color depending on whether color Doppler imaging is visible in the video clip.
- ii. Frame Quality*
 - o Each frame will be rated for quality on a three-tier scale: Poor, Normal, or Good. This parameter will help filter and prioritize frames for training.
- iii. Metadata Integration*
 - o Labels will be linked to metadata, including patient-specific information (age, sex, diagnosis) and labeling-specific attributes (labeling cardiologist ID, timestamp).

b. Labeling Workload and Quality Assurance

Because the labeling process is demanding, each video clip will be labeled by a single cardiologist. To ensure inter-rater reliability, 5% of each cardiologist's labeled data will be randomly assigned to a second cardiologist for cross-validation and variability assessment.

We anticipate the participation of 10 cardiologists, each assigned approximately 500 video clips to label over a period of 3 months.

2. Prospective Data Labeling: Automated Label Derivation

The prospective data collected during routine echocardiography scans already includes labels of the cardiac views assigned by the cardiologist or sonographer. However, these labels cannot be directly used to train the AI model, as the model would simply learn to identify the labels on the frames instead of understanding the actual cardiac views. This would limit the model's ability to generalize and recognize unlabelled cardiac views in real-world applications.

To overcome this challenge, the team will develop an innovative automated script to extract unlabelled cardiac views from the labelled prospective data. This script will analyze the ultrasound video clips, identifying pauses that indicate when the expert saved specific views. It will then determine the names of these saved views based on the expert-assigned labels. Finally, the script will extract unlabelled video segments from between these pauses, which represent the same cardiac view labelled by the expert but without visible labels.

The unlabelled clips, along with their associated labels stored separately, will then be uploaded into Labelbox for further labeling of video quality and any additional view-specific parameters. The goal is to provide diverse examples of the same cardiac view for model training while keeping the labels hidden during the training phase.

Metadata Extraction and Format

After labeling, data will be exported in JSON format from the Labelbox platform and processed for model training in the Python environment provided by the Ilifu platform. The metadata will adhere to a predefined structure to ensure consistency and usability across the dataset. The format will include the following key details: Video ID, Cardiac View, Doppler Color, Frame Quality, Labeling Cardiologist ID, Labeling Timestamp, Is Labelled, and Label Associated Field. (Figure 3) below illustrates the standardized format for labeled video metadata, ensuring consistency and usability for model training. This structured format will ensure that labeled and unlabelled data are handled separately while remaining consistent in terms of metadata attributes. This approach will facilitate tracking and integrating the data into the model training pipeline.

Figure 3: Sample of Labelled Video Clip Metadata Structure in JSON format.

Every labeled video exported from Labelbox will follow this predefined format, ensuring uniformity and ease of integration. Each video will be uniquely identified by a "Video ID," such as CHDAIP0000101. The labels of interest extracted from the labelled dataset include the Cardiac

Views, the Frame Quality and the Doppler Color. The "Cardiac View" parameter will specify the view of the video, with the available options detailed in (Table 1). The "Doppler Color" parameter will indicate whether color Doppler imaging is present, labeled as either "Color" or "No Color." The overall quality of the video frames will be assessed under the "Frame Quality" parameter, categorized as "Good," "Medium," or "Poor" based on clarity and diagnostic usability. Each video will also be associated with a "Labeling Cardiologist ID," which will serve as a unique identifier for the cardiologist responsible for labeling the video, along with a "Labeling Timestamp" to record the exact date and time the labeling is completed. The "Is Labelled" parameter will differentiate between frames that are prospectively labeled and those extracted by an automated script. Finally, for unlabelled clips, a "Label Associated Field" will contain labels derived from the automated script, ensuring these clips can still be utilized in analysis.

This structured format will not only ensure consistency in metadata attributes but will also enable efficient tracking of labeled and unlabelled data. It will streamline the integration of these data points into the model training pipeline, improving the overall accuracy and reliability of the process. The flowchart in (Figure 4) below depicts the entire process.

Figure 4: Metadata Processing Steps

Development of the AI model

The proposed AI model will benefit from the successful and validated works of Mason et al., (2019) [18] and Madani et al., (2018) [19]. The development of the model will involve four key steps: data collection and preparation, data preprocessing, data exploration and visualization, model building and model evaluation.

1) Data preprocessing

Data processing will involve two major tasks: 1) improvement of pictorial information for human interpretation and 2) processing of image data for storage, transmission and representation for autonomous machine perception. The data preprocessing steps can be broken into low-, mid- and high-level processes.

- i. **Low-Level Processes:** Images and image frames extracted from videos will undergo primitive operations using image processing techniques such as Gaussian filtering for noise reduction, histogram equalization for contrast enhancement, and unsharp masking for image sharpening. These steps ensure improved clarity and quality of input images before further processing.
- ii. **Mid-Level Processes:** Following the procedure of Mason et al., (2019), [18] images from the low-level processing step above will be further processed by:
 - o removing doppler and split views (Segmentation) and
 - o sampling random background frames.
- iii. **High-Level Processes:** Here images from mid-level processing are further processed by:
 - o cropping field of view and rescaling,
 - o Inpainting labels and annotations of the images, and
 - o splitting the data into training and test sets.

2) Data exploration and visualization

This step will involve looking for common data problems when training the models, some of which may fall under object detection, since we will be detecting the cardiac

plane and the heart structure. We anticipate problems related to:

- o image dimensions and aspect ratios (i.e. images having different sizes, for example: one image can be of size 32*32 and another can be of size 123*47), and
- o label composition – imbalances, bounding box sizes, aspect ratios. (i.e. some cardiac views may have more images compared to others).

To address these issues, we will use dedicated tooling like google facets (a data visualization tool) to explore image data and take decisions accordingly.

3) Model Building

a. Phase 1 model development

For detecting cardiac views and localizing heart structures in newborns using real-time echocardiography videos, we will adopt a novel approach inspired by Mason et al. (2019) [18]. Their work demonstrated an efficient pipeline for video object detection, achieving state-of-the-art performance.

To adapt their architecture for our dataset, we propose an interleaved framework with two feature extractors of distinct speeds and recognition capabilities. These extractors will process different frames, and the extracted features will maintain a common visual memory through a convolutional Long Short-Term Memory (convLSTM) layer, which fuses context from prior frames with the gist from the current frame to generate detections.

Following Mason et al.'s approach, the interleaving policy determining when to run each feature extractor will be learned through reinforcement learning. The convLSTM is particularly suited for sequential image data, making it ideal for analyzing short echocardiography video clips.

Training Specifics

- i. **Dataset Composition:** The dataset comprises short echocardiography video clips in AVI (St. Padre Pio Clinic, Cameroon) and DICOM formats (UCT, South Africa), along with metadata files (Excel format) containing age, sex, and diagnosis.
- ii. **Metadata Utilization:** Metadata will be incorporated into the model's training to enhance feature representation, allowing the model to exploit diagnostic context alongside visual data.
- iii. **Labeling Annotations:** Video clips will be annotated on the Labelbox platform with:
 - o **Cardiac view** (10 predefined categories, with sub-parameters for Doppler presence).
 - o **Video quality** (poor, normal, or good). Exported JSON annotations will be integrated into the training pipeline.

Figure 5 An illustration of a memory-guided interleaved model (Mason et al., 2019)

11 01

From (Figure 5), given a video stream, a visual memory model fuses diverse visual features from fast and slow feature extractors across frames to generate live detections. The layers and weights of this pretrained model will be fine-tuned on our dataset, focusing on the specific views of the heart in newborns.

Training Process

. The training sessions will follow an iterative approach, with checkpoints saved at regular intervals to ensure seamless resumption in case of interruptions. The model will first be trained from scratch using the collected and constructed dataset. As additional data becomes available from various sources, a fine-tuning approach will be employed to gradually enhance performance and adapt the model to new variations. The training will be conducted using libraries such as TensorFlow and PyTorch, leveraging high-performance GPUs available on the computing cluster to optimize training efficiency and scalability.

b. Phase 2: Model refinement

In Phase 2, prospective data will be used to refine the initial model, ensuring enhanced accuracy and adaptability. The model will leverage prediction confidence (PC) to optimize video clip extraction and quality assessment:

i. Dynamic Prediction Confidence (PC):

Given an ultrasound video stream, the model will compute the probability (PC) that an extracted image frame or video clip belongs to a specific cardiac view.

- o Video clip extraction will begin once a defined PC threshold is met on an image frame.
- o Subsequent frames that meet the PC threshold will be compiled into a 5-second video clip for the identified view.
- o In cases where multiple video clips of the same view are generated, the clip with the highest average PC across image frames will be retained as the final clip.

ii. Video Quality Optimization:

Since image quality can vary depending on the ultrasound device, the model will enable users to set a PC threshold at first use to ensure the best image quality:

- o **High PC:** If the initial PC setting is higher than the device's capabilities, the model will prompt the user to lower the PC threshold.
- o **Low PC:** Conversely, if the initial PC setting is too low, the model will suggest increasing the threshold to maximize quality.

These mechanisms will ensure that the best possible images are extracted

regardless of device variability. An associated manual will guide users to begin with higher PC settings during initial model use.

iii. Integration with Training Pipeline:

- o During model refinement, metadata (e.g., age, sex, diagnosis) will be used alongside visual data to enhance feature representation and improve diagnostic accuracy.
- o Labels generated during the annotation phase (e.g., cardiac view, Doppler presence, and video quality) will be incorporated into the refinement process.

(Figures 6) and (Figure 7) illustrate the sequence and flow of the video clip extraction and PC adjustment mechanisms during real-time ultrasound scans.

Figure 6 Sequence diagram of our AI model system

IV. Refinement Goals:

- o To integrate metadata-driven learning, leveraging age, sex, and diagnosis to fine-tune the model.
- o To enhance adaptability for various ultrasound devices while ensuring robust performance.

V. Expected Outcomes

The proposed protocol combines state-of-the-art video object detection with metadata integration, enabling robust detection of cardiac views and heart structure localization in echocardiography video streams. The model's reinforcement learning-based interleaving policies, dynamic PC adjustments, and quality control mechanisms ensure adaptability across devices and datasets.

Figure 7 (a). Extract of sequence diagram to show how the highest frame accuracy will be obtained.

Figure 7 (b). Extract of sequence diagram to show how the video clips for required view will be obtained.

4) Model Evaluation

To comprehensively assess the model's performance in detecting cardiac views and localizing heart structures in neonates, a robust evaluation framework will be implemented. The evaluation will follow an 80/20 training-testing split, ensuring a representative distribution of data. Additionally, k-fold cross-validation (e.g., 10-fold cross-validation) will be used to improve the reliability of performance metrics by averaging results across multiple splits. The evaluation metrics will include:

- i. **Overall Accuracy:** The percentage of correctly classified instances across all

classes in the test dataset.

- ii. **Per-Class Accuracy:** The accuracy for each individual cardiac view, providing insights into the model's performance across different heart structures.
- iii. **Average Accuracy:** The mean accuracy across all classes, offering a balanced assessment of class-specific predictions.
- iv. **Confusion Matrix Analysis:** Confusion matrices will be generated to visualize true positives, false positives, true negatives, and false negatives for each class. These matrices will help identify systematic misclassifications and underperforming classes.
- v. **F1-Score:** The harmonic mean of precision and recall, calculated for each class, to measure the balance between false positives and false negatives. The macro F1-score will summarize the overall classification performance.
- vi. **Receiver Operating Characteristic (ROC) Curve and Area Under the Curve (AUC):** For binary classification tasks (e.g., Doppler vs. non-Doppler views), ROC curves will be plotted to illustrate the trade-off between sensitivity and specificity. The AUC will quantify the model's discriminatory ability.
- vii. **C-Statistics:** Calculated to measure the concordance between predicted and actual classes, serving as an overall indicator of the model's diagnostic capability.
- viii. **Saliency Maps:** Visualization techniques will be applied to highlight regions in input images that strongly influence the model's predictions, enhancing interpretability and ensuring alignment with domain knowledge, such as focusing on heart structures in echocardiography images. To support explainable AI, methods such as Grad-CAM, Grad-CAM++, and other class activation mapping techniques will be utilized. These approaches will help visualize the model's decision-making process, providing deeper insights into how specific features contribute to classification outcomes.
- ix. **Class-Specific Evaluation Metrics:**
 - o **Precision** (positive predictive value): The fraction of correctly predicted positive observations.
 - o **Recall** (sensitivity): The ability of the model to identify all relevant instances in a class.
 - o **Specificity:** The ability of the model to exclude irrelevant instances.
- x. **Video-Based Metrics:** For evaluating performance on the short video clips:
 - o **Clip Prediction Accuracy:** The percentage of correctly classified video clips, where the assigned view corresponds to the majority or highest-confidence prediction across frames.
 - o **Temporal Consistency:** The consistency of predictions within a video clip, measured as the proportion of consecutive frames with matching predictions.
- xi. **Metadata-Driven Analysis:** Using age, sex, and diagnosis data from the metadata, the performance will be evaluated for specific subpopulations (e.g., neonates with congenital heart defects vs. those without). This ensures that the model performs robustly across diverse patient demographics and conditions.

By combining these quantitative and qualitative evaluation metrics, the model's accuracy, reliability, and interpretability will be rigorously assessed. The evaluation will follow the approach of Madani et al., (2018) [19] but will be extended to include video-based and metadata-specific metrics to account for the unique characteristics of the dataset

Workflow

A GitHub repository for the project will be created and different members of the engineering group will have access to this repository so that they can clone it to their local machines to make contributions. A branch protection rule will also be created so that the code contributed by engineers is reviewed by the admin (Senior engineer). The repository will contain the work of Mason et al, 2019 freely available on GitHub.

SOP and instructional video on echocardiography sweep for non-experts

Making use of relevant materials from the literature, our expert cardiologist in the research team will develop a step-by-step procedure on how to conduct an echocardiography sweep that will allow the AI model to capture the defined frames/cardiac views of interest. The SOP shall detail preparatory steps (set up) prior to initiating an echo scan, tips on probe handling and manoeuvre, the sequence of moving the probe across the baby's chest, including start and end positions. Once the SOP is finalized, this will be used to develop an instructional video to facilitate training of the non-experts.

Investigating the diagnostic utility of the video clips produced from the AI model in relation to their use in clinical practice

We will investigate the usefulness of the AI model extracted video clips in clinical practice through the following steps.

- a. **Develop a platform to operate the model.** The platform will be available on a computer (local server) and will act as an interphase for the AI model to process the uploaded live video stream.
- b. **Obtain livestream echocardiography videos from clinical practice.** We will recruit 4 non-experts (GP, nurse, midwife and sonographer) who will receive basic training and guidance on postnatal echocardiography and basic probe sweep using the SOP and instructional videos developed. Practice sessions will take place at St Padre Pio Hospital in Douala-Cameroon. After the training, the non-experts will be asked to perform test echocardiography scan (basic sweep) for neonates referred to St Padre Pio Hospital in Douala-Cameroon to obtain video streams. The test scan will take place after the baby's normal scan procedure has been completed and we expect each session to last for about 3 minutes (3 minutes video clip). The entire exercise is expected to last for four months, during which at least 16 neonates are expected to have been scanned (16 video clips obtained). It is important to emphasize that the scan from the non-experts are for study purposes only and will not replace the normal specialist scan for the baby. However, if found relevant after evaluation, the images/video clips produced by the AI model could be used by the expert to inform any diagnosis. For each baby,

the scan video by the experts will be retained alongside the test videos from the non-experts. All videos will be deidentified using our anonymization script. All parents will receive orientation to the exercise and provide signed informed consent, with the choice not to consent (see appendix C for information sheet and consent form).

- c. **Extraction and labelling of images/video clips from the ultrasound videos.** The videos from the expert and non-expert ultrasound sessions will be loaded into the AI model platform in the computer for the cardiac views (short video clips) to be extracted and sent for expert review.
- d. **Evaluate images/video clips extracted.** The video clips will be randomly distributed to our panel of specialist paediatric cardiologist for independent evaluation on the quality of the extracted video clips/frames, correctness of the AI assigned frame label, and usability of the video clips generated by the AI model in real practice based on a 5-point Likert scale. Additionally, the experts will examine the images for any CHD diagnosis as they would in clinical practice. At the end of the exercise the results will be tallied and the panellists reconvened, chaired by one of the members as shall be determined, to deliberate on results where there are disagreements. A comparison will also be made between the video clips obtained by the expert and non-experts. The review process is expected to last for 3 months.

Ethical considerations and data security

In developing this methodology, we have paid particular attention to ethical considerations that will ensure we build an AI model that is not only effective and efficient but also equitable, transparent, and respectful of patient rights and cultural contexts. We aim to balance technological innovation with patient safety, data privacy, and the equitable distribution of benefits. Accordingly, key ethical aspects have been addressed to ensure compliance with international standards and local regulations.

Given the focus on neonatal research, we have made informed consent a cornerstone of the study. Parents or legal guardians of participating neonates will be fully informed about the study's objectives, procedures, potential benefits, and risks. This process ensures that participation is entirely voluntary and that parents understand how their child's data will be used. Consent forms will be provided in appropriate local languages, and researchers will be available to answer any questions, thereby addressing potential literacy barriers and fostering trust.

Neonatal echocardiography involves sensitive medical data, and maintaining its confidentiality is essential. All imaging data and associated patient information will be anonymized before storage or analysis. Encryption protocols and secure data transfer methods will be used to prevent unauthorized access. These measures comply with international guidelines such as the General Data Protection Regulation (GDPR) and national data protection laws of study countries.

We recognise the fact that equitable access to the benefits of the AI model is a critical ethical priority. SSA faces significant disparities in healthcare access, and introducing this model could inadvertently exacerbate existing inequalities if not implemented thoughtfully. By design, this study will prioritize deploying the AI tool in low-resource settings where access to expert cardiologists is limited, ensuring that the intervention aligns with its goal of addressing healthcare inequities.

Another major ethical concern is the potential for bias in AI algorithms. The model's training data must represent the diverse populations it aims to serve to ensure diagnostic accuracy across different demographic groups. Inadequate representation could lead to disparities in model performance, undermining trust and fairness. To mitigate this, the study will incorporate data from diverse neonatal populations in Cameroon and South Africa and plans to expand the dataset to include additional low-resource settings.

We are keen on transparency and accountability which are integral to ethical the deployment of AI tools. Healthcare providers using the AI tool must understand its capabilities and limitations. We are committed to ensuring that the AI tool does not replace clinical judgment but rather serve as a supportive aid for non-experts. We will implement clear communication about the tool's accuracy, error rates, and scope of applicability in order to prevent over-reliance and promote responsible use. Additionally, the study will establish mechanisms for continuous monitoring and reporting of any adverse events or unexpected outcomes during the implementation phase.

Lastly, we recognise the fact that the long-term sustainability and scalability of the AI model may raise ethical questions about resource allocation and stakeholder involvement. We aim to engage local healthcare professionals, policymakers, and community leaders early in the process to ensure that the model is culturally appropriate and addresses the specific needs of the target population. We have incorporated measures to build local capacity of the non-experts on how to use and maintain the AI tool thereby preventing dependency on external entities and fostering self-reliance in healthcare system.

Results

Retrospective data collection for the project began in September 2024, and since then, data from 308 babies have been collected. Of these, data from 145 babies in Cameroon have already been labelled, while data from 163 babies in South Africa is still being transmitted. In parallel, the project model is under development, with the labelled data from Cameroon being used to initiate training. Although the initial test results are not optimal, given that all the project data must be labelled and consolidated into a single dataset, these preliminary tests serve as a calibration step. Overall, the data collection process remains ongoing, and we anticipate that full data collection will be completed by May 2025. Model completion,

validation and publication of the complete results will follow thereafter.

Discussion

Modern advancements in data science and digital health technologies present unique opportunities to transform healthcare delivery in low-resource settings. This protocol outlines a structured approach to developing an AI-assisted echocardiography model aimed at empowering non-experts to perform neonatal cardiac scans and obtain accurate images for expert review. By addressing critical gaps in CHD diagnosis, the project takes an important step toward improving early detection and referral pathways, ultimately enhancing survival rates for neonates in SSA.

A key next step would be the integration of the AI model with existing pulse oximetry and clinical ultrasound systems, and the establishment of a telemedicine platform to facilitate image transmission to remote paediatric cardiologists. One potential platform for such integration is the Commonwealth Award-winning Global Birth Defects (GBD) app [20], designed to facilitate the accurate diagnosis of CAs for research and surveillance programs operating in low resource settings. The app already has the option to collect data on CAs (including images and videos) and send to an expert panel for evaluation. Currently the app is mainly designed for externally visible CAs. The inclusion of assistance with pulse oximetry and echocardiograms via the CHD AI model would make the app more comprehensive, and could transform it beyond surveillance and research to a clinical tool. This possibility also falls within the remit of our new sub-Saharan African Congenital Anomaly Network (sSCAN - [21]) which can then lead the effort for wide scale adoption within SSA and other LMICs.

The adoption of AI in clinical practice is not without challenges, including skepticism, ethical concerns, and regulatory requirements. To address these, the protocol incorporates processes for rigorous clinical validation and stakeholder feedback. This includes piloting the AI model with target users, evaluating the quality of AI-extracted images by expert paediatric cardiologists, and gathering healthcare providers' opinions on the feasibility of AI integration into routine practice. These steps ensure alignment with ethical standards and enhance the model's clinical relevance and acceptance. As the current AI model is limited to image capture and does not suggest a CHD diagnosis, it is expected to promote greater acceptance among healthcare providers and the public.

The inclusion of diverse patient populations from Cameroon and South Africa further strengthens the external validity of the proposed model. This diversity more effectively captures regional differences in healthcare infrastructure and patient demographics, improving its broader applicability across SSA. However, challenges in generalizing AI models to diverse contexts remain a significant barrier [22]. Expanding the dataset to include additional SSA countries and other low-resource settings will enhance the model's robustness and facilitate wider adoption.

The protocol's imaging sequences and cardiac views adhere to the 2024 American Society of Echocardiography (ASE) guidelines [16], ensuring a standardized approach to data collection. Additionally, the dual-phase model development strategy—leveraging both retrospective and prospective datasets—addresses limitations in data availability while adhering to ethical guidelines. The incorporation of metadata-driven insights and quality assurance protocols further strengthens the model's diagnostic accuracy and clinical utility.

Future updates to the current model could provide real-time feedback to guide users in probe maneuvers, improving both image accuracy and user expertise. As the model evolves, its scope can expand to include older paediatric populations, prenatal echocardiography, automated CHD subtype detection, and predictive analytics for treatment planning and survival outcomes. These advancements could enhance the robustness and value of the AI tool, redefining and broadening its impact on CHD management in resource-limited settings.

In conclusion, this protocol exemplifies a rigorous, ethical, and scalable approach to leveraging AI for neonatal echocardiography in low-resource settings. The unique approach of integrating the AI model with established clinical and research standards will set the stage for broader development and adoption of AI technologies to improve neonatal health outcomes in SSA and other low resource settings.

Acknowledgements

We acknowledge the National Institutes of Health (grant award number 1U01HL172179-01). The content of this article is solely our responsibility and does not necessarily represent the official views of the US National Institutes of Health.

Liesl Zühlke is funded by the South African Medical Research Council (SAMRC) through its Division of Research Capacity Development under the Mid-Career Scientist Programme from funding received from the South African National Treasury. The content and findings reported/illustrated are the sole deduction, view and responsibility of the researcher and do not reflect the official position and sentiments of the SAMRC.

Conflicts of Interest

The authors declare no competing interests.

Abbreviations

AI: Artificial Intelligence
ASE: American Society of Echocardiography
AUC: Area Under the Curve
AVI: Audio Video Interleave
AWS: Amazon Web Services
CA: Congenital Anomaly
CHD: Congenital Heart Disease
GBD: Global Birth Defects
convLSTM : convolutional Long Short-Term Memory
DICOM: Digital Imaging and Communications in Medicine
DSI: Data Science for Health Discovery and Innovation
GCP: Google Cloud Platform
GDPR: General Data Protection Regulation
GPU: Graphical Processing Unit
IQR: Interquartile Range
JSON: JavaScript Object Notation
LMICs: Low- and Middle-Income Countries
ODSP: Open Data Science Platform
PC: Prediction Confidence
PII: Personally Identifiable Information
ROC: Receiver Operating Characteristic Curve
SOP: Standard Operating Procedure
SSA: Sub-Saharan Africa
sSCAN: sub-Saharan African Congenital Anomaly Network

References

1. UNIGME-2024-Child-Mortality-Report, United Nations Children's Fund (UNICEF), 2025, ISBN: 978-92-806-5636-7
2. Su, Z., Zou, Z., Hay, S. I., Liu, Y., Li, S., Chen, H., ... & Zhang, H. (2022). Global, regional, and national time trends in mortality for congenital heart disease, 1990–2019: An age-period-cohort analysis for the Global Burden of Disease 2019 study. *EClinicalMedicine*, 43. Jivanji SG, Lubega S, Reel B, Qureshi SA. Congenital heart disease in East Africa. *Frontiers in pediatrics*. 2019 Jun 26;7:250.
3. landscape analysis of paediatric cardiac services in Africa, joint publication: *Cardiology In The Young* and *World Journal for Pediatric and Congenital Heart Surgery*, 2025: *Cardiology In The Young* and *World Journal for Pediatric and Congenital Heart Surgery*, 2025
4. Saxena, A., 2019. Status of pediatric cardiac care in developing countries. *Children*, 6(2), p.34.

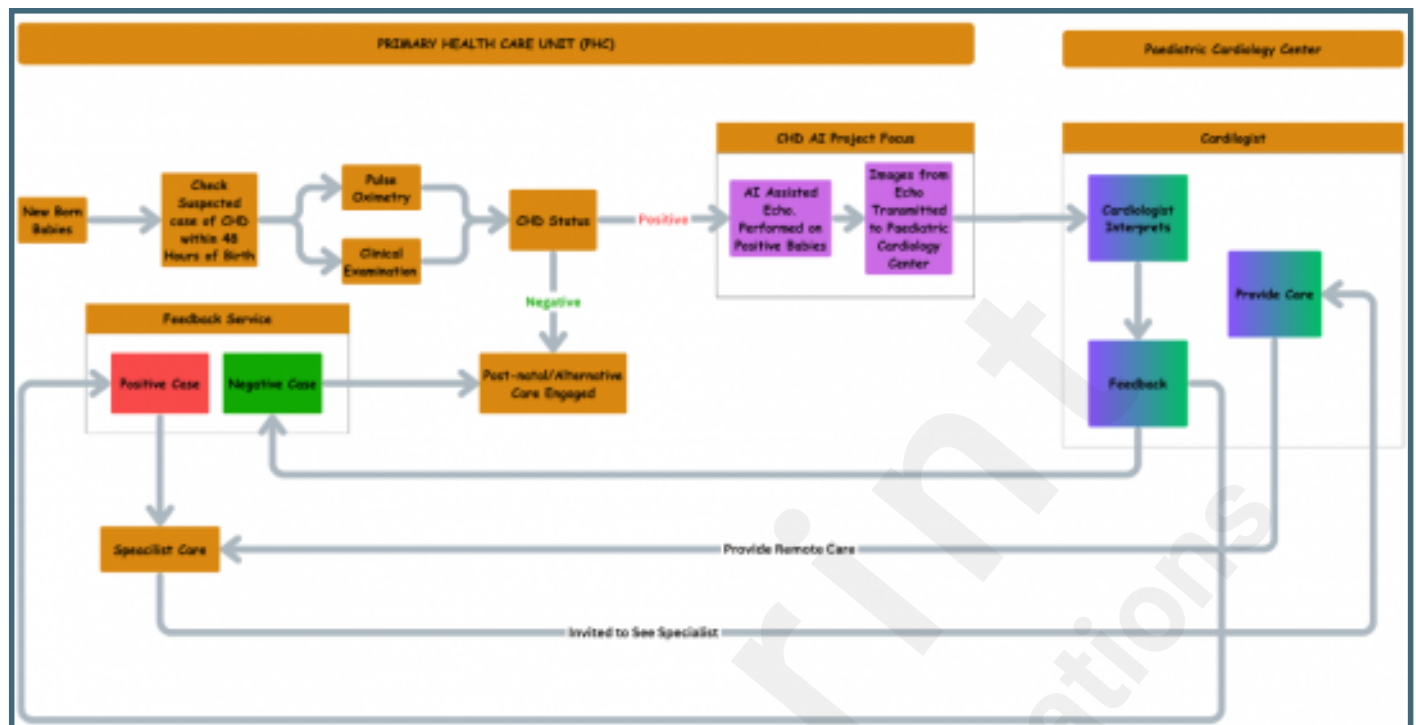
5. Iyer, P.U., Moreno, G.E., Caneo, L.F., Faiz, T., Shekerdemian, L.S. and Iyer, K.S., 2017. Management of late presentation congenital heart disease. *Cardiology in the Young*, 27(S6), pp.S31-S39.
6. Van Niekerk, A. M., Cullis, R. M., Linley, L. L., & Zühlke, L. (2016). Feasibility of Pulse Oximetry Pre-discharge Screening Implementation for detecting Critical Congenital heart Lesions in newborns in a secondary-level maternity hospital in the Western Cape, South Africa: The 'POPSICLe'study. *South African Medical Journal*, 106(8), 817-820.
7. Cabanas AM, Fuentes-Guajardo M, Latorre K, León D, Martín-Escudero P. Skin Pigmentation Influence on Pulse Oximetry Accuracy: A Systematic Review and Bibliometric Analysis. *Sensors*. 2022 Apr 29;22(9):3402.
8. Vesoulis Z, Tims A, Lodhi H, Lalos N, Whitehead H. Racial discrepancy in pulse oximeter accuracy in preterm infants. *Journal of Perinatology*. 2022 Jan;42(1):79-85.
9. Murni IK, Wibowo T, Arafuri N, Oktaria V, Dinarti LK, Panditwa D, Patmasari L, Noormanto N, Nugroho S. Feasibility of screening for critical congenital heart disease using pulse oximetry in Indonesia. *BMC pediatrics*. 2022 Dec;22(1):1-8
10. Goley SM, Sakula-Barry S, Kelly A, Wright N. Investigating the use of ultrasonography for the antenatal diagnosis of structural congenital anomalies in low-income and middle-income countries: systematic review protocol. *BMJ paediatrics open*. 2019;3(1).
11. Day TG, Kainz B, Hajnal J, Razavi R, Simpson JM. Artificial intelligence, fetal echocardiography, and congenital heart disease. *Prenatal Diagnosis*. 2021 May;41(6):733-42.
12. Tromp, J., Seekings, P. J., Hung, C. L., Iversen, M. B., Frost, M. J., Ouwerkerk, W., ... & 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.
13. Wehbe, R. M., Katsaggelos, A. K., Hammond, K. J., Hong, H., Ahmad, F. S., Ouyang, D., ... & Thomas, J. D. (2023). Deep learning for cardiovascular imaging: A review. *JAMA cardiology*, 8(11), 1089-1098.
14. Ueda, D., Matsumoto, T., Ehara, S., Yamamoto, A., Walston, S. L., Ito, A., ... & Miki, Y. (2023). Artificial intelligence-based model to classify cardiac functions from chest radiographs: a multi-institutional, retrospective model development and validation study. *The Lancet Digital Health*, 5(8), e525-e533.
15. Christensen, M., Vukadinovic, M., Yuan, N., & Ouyang, D. (2024). Vision–language foundation model for echocardiogram interpretation. *Nature Medicine*, 30(5), 1481-1488.
16. Mcnamara, Patrick & Jain, Amish & EL-Khuffash, Afif & Giesinger, Regan & Weisz, Dany & Freud, Lindsey & Levy, Philip & Bhombal, Shazia & Boode, Willem-Pieter & Leone, Tina & Richards, Bernadette & Singh, Yogen & Acevedo, Jennifer & Simpson, John & Noori, Shahab & Lai, Wyman. Guidelines and Recommendations for Targeted Neonatal Echocardiography and Cardiac Point-of-Care Ultrasound in the Neonatal Intensive Care Unit: An Update from the American Society of Echocardiography. *Journal of the American Society of Echocardiography* (2024). 37. 171-215. 10.1016/j.echo.2023.11.016.
17. Gearhart A, Goto S, Deo RC, Powell AJ. An automated view classification model for pediatric echocardiography using artificial intelligence. *J Am Soc Echocardiogr*.

- (2022) 35(12):1238–46. 10.1016/j.echo.2022.08.009 [PMC free article] [PubMed] [CrossRef] [Google Scholar]
18. Mason L, Menglong Z, Marie W, et al. Looking fast and slow: Memory guided mobile video object detection. arxiv. 2019;1903.10172v1
 19. Madani, A., Arnaout, R., Mofrad, M. and Arnaout, R., Fast and accurate view classification of echocardiograms using deep learning. npj Digit. Med. 1, 6 (2018).
 20. Dolk, H., Leke, A.Z., Whitfield, P., Moore, R., Karnell, K., Barišić, I., Barlow-Mosha, L., Botto, L.D., Garne, E., Guatibonza, P. and Godfred-Cato, S., 2021. Global birth defects app: An innovative tool for describing and coding congenital anomalies at birth in low resource settings. *Birth defects research*, 113(14), pp.1057-1073.
 21. Sub-Saharan Congenital Anomalies Network. (n.d.). The Global Health Network. <https://sscan.tghn.org/>
 22. Dangi, Ravi & Sharma, Anil & Vageriya, Vipin. Transforming Healthcare in Low-Resource Settings With Artificial Intelligence: Recent Developments and Outcomes. Public health nursing (Boston, Mass.) (2024). 10.1111/phn.13500.

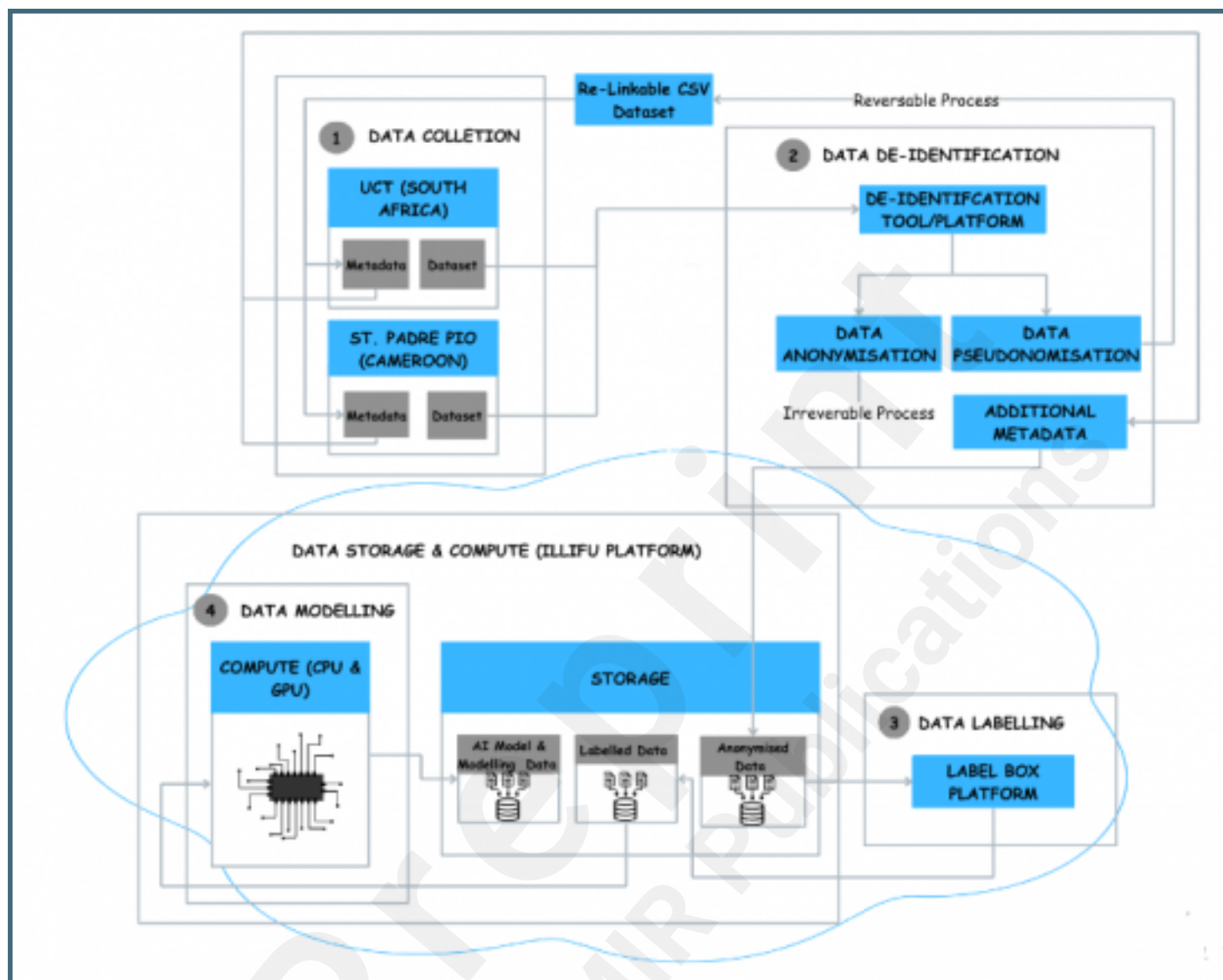
Supplementary Files

Figures

AI assisted echocardiography CHD screening process.



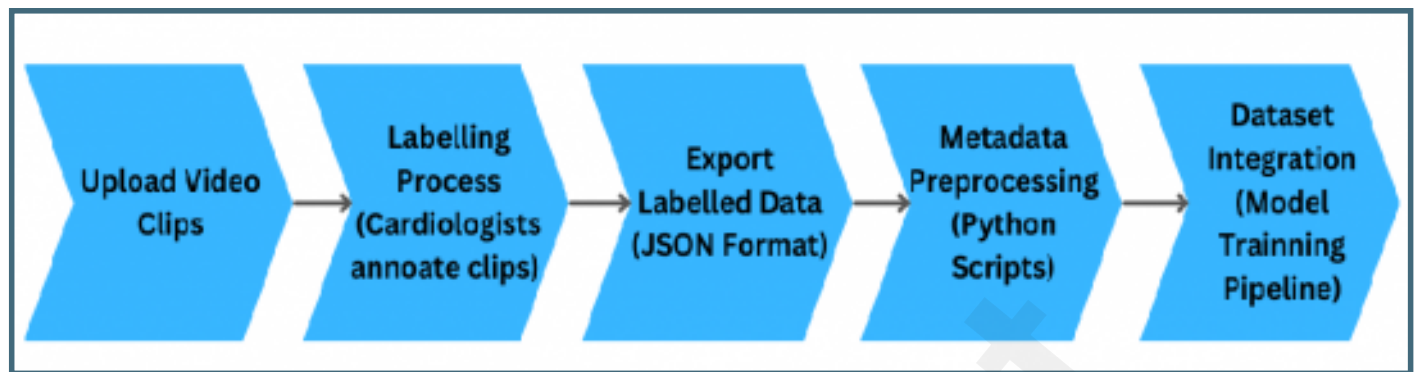
Thematic data flow diagram for CHD-AI project.



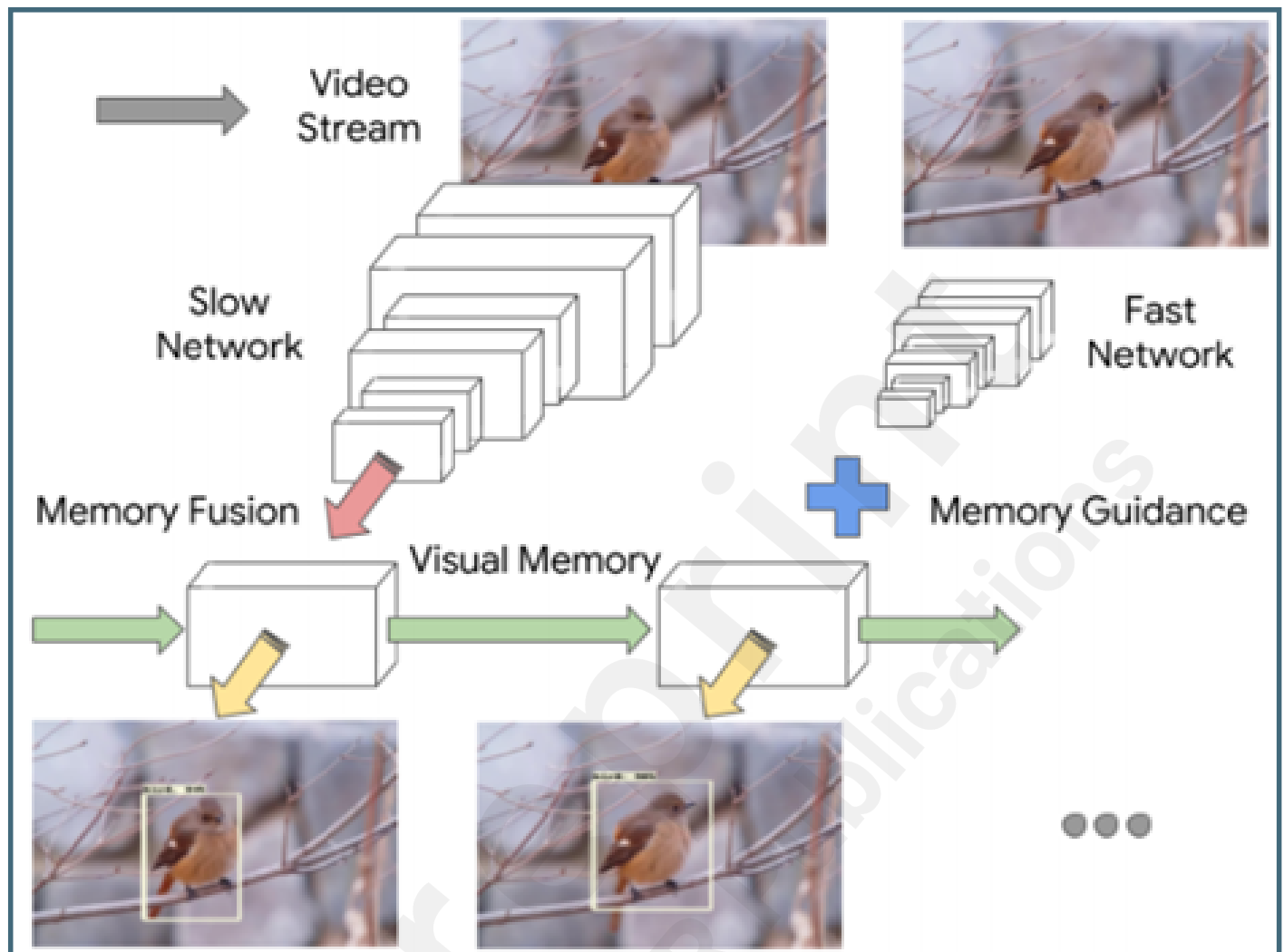
Sample of Labelled Video Clip Metadata Structure in JSON format.

```
{  
  "VideoID": "CHDAIP00000101",  
  "CardiacView": "Apical Four-Chamber",  
  "DopplerColor": "Color",  
  "FrameQuality": "Good",  
  "LabelingCardiologistID": "CARD001",  
  "LabelingTimestamp": "2025-01-02T14:35:00Z",  
  "IsLabelled": true,  
  "LabelAssociatedField": "Normal Function"  
}
```

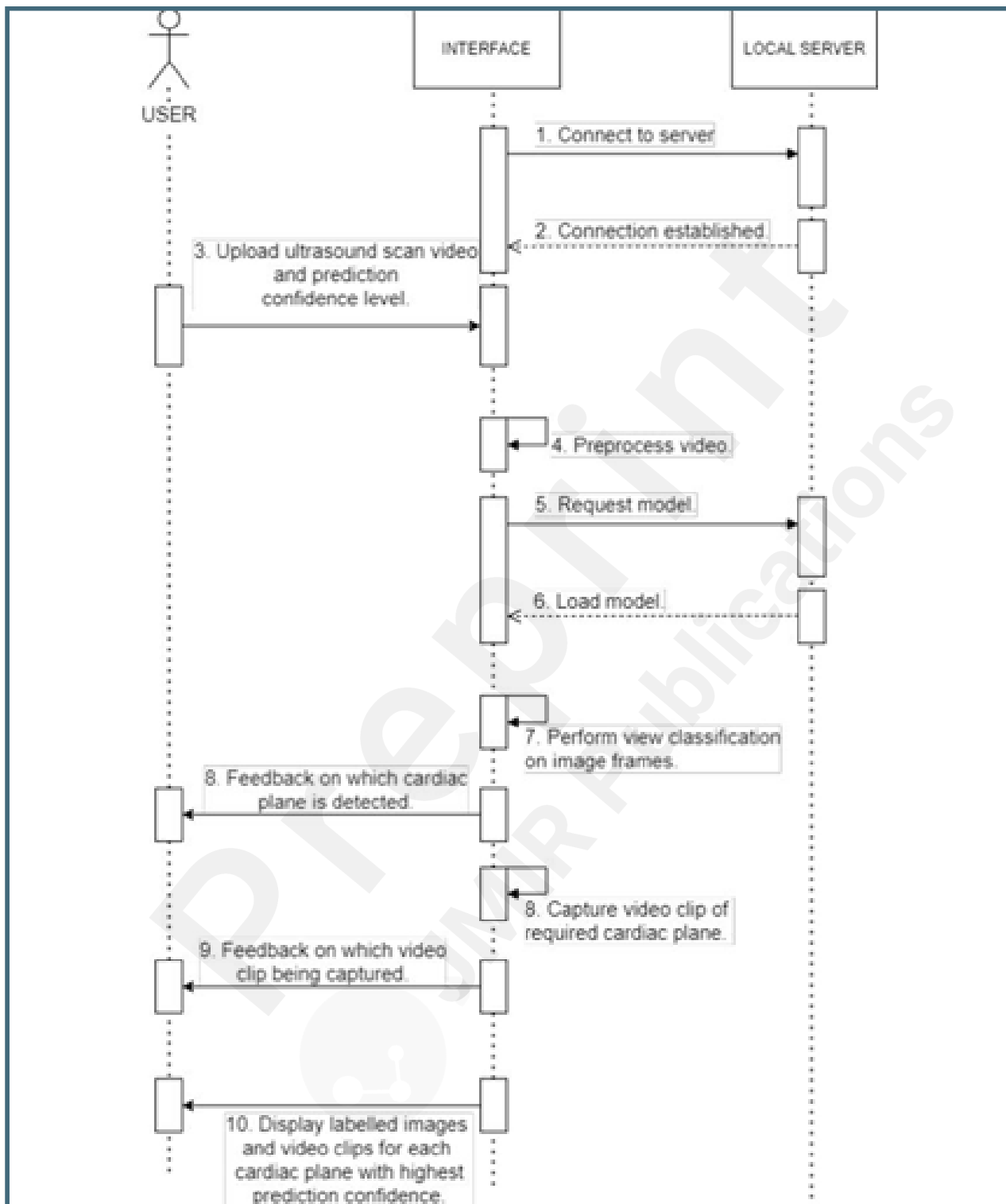
Metadata Processing Steps.



An illustration of a memory-guided interleaved model.



Sequence diagram of our AI model system.



Extractions of sequence diagram.

