The Role of Artificial Intelligence in Streamlining Echocardiography Quantification

White Paper

–

Kristin McLeod
Data Science Leader,
Cardiovascular Ultrasound

Vivid

The Role of Artificial Intelligence in Streamlining Echocardiography Quantification

Introduction

The burden of tedious tasks as demand for echocardiograms continues to grow

Ischemic heart disease and stroke are leading causes of death, accounting for a combined 17.9 million deaths in 2016, and have remained the leading causes of death globally in the last 15 years.[1] By 2030, 40.5% of the US population is projected to have some form of Cardiovascular Disease (CVD).[2]

Echocardiography has become standard in the diagnosis, management, and follow-up of patients with any suspected or known heart diseases and is one of the most widely used diagnostic tests in cardiology. Considering the ever-growing demand for echocardiograms, cardiologists and sonographers often spend valuable time adjusting parameters and repeating tests to obtain the same measurements. These repetitive tasks require the consideration of numerous parameters, and multiple button clicks, just to input measures. A study indicated 90% of clinical sonographers experienced symptoms of WRMSDs (Work Related Musculoskeletal Disorders).[3]

Staff shortages due to injuries and increasing referrals for sonography have resulted in insufficient rest periods, further increasing the duration of the sonographer's exposure to risk and generating up to $120+ billion yearly in direct and indirect costs for employers.[2]

90%
of sonographers experienced Work Related Musculoskeletal Disorders

$120+ billion
yearly in direct and indirect costs for employers

Poor quality has a cost, standardization drives improvement

Cardiac ultrasound can provide important information in critical and emergency settings that help users save lives. Data acquisition depends on specific imaging targets, conditions or scenarios, the ultrasound equipment used, techniques and protocols applied, and related to the level of training and skill of the operator and the individual operator's profile. Making the task even more difficult, other variables like patient size and echogenicity, may impact overall image quality, measurement accuracy and variations between operators.[4]

Although historically cardiologists were almost exclusively responsible for performing, supervising and interpreting echocardiographic examinations in acute and emergency settings, fully trained cardiologists are not always available where medical emergencies occur.[5]

Standardization of structure and process increase the likelihood of desired health outcomes.[6]

AI can help refocus clinical teams on core strengths and reduce inter-operator dependency

As detailed in a recent briefing from McKinsey, Artificial intelligence (AI) has the potential to transform how care is delivered.[7]

This improved efficiency allows healthcare systems to provide better care to more people and can help improve the experience of healthcare practitioners, enabling them to spend more time in direct patient care and reducing burnout.

At GE Healthcare, we believe that AI can support the faster delivery of care, by enabling accelerated diagnosis time, and help healthcare systems manage population health more proactively, allocating resources to where they can have the largest impact. With the Vivid™ Ultra Edition, we bring AI capabilities to the entire portfolio so every institution can benefit from these virtual assistants in daily practice.

Ultra Edition brings AI to the entire Vivid portfolio

AI is now playing a critical role

2D (B-mode) imaging and spectral Doppler imaging remain the cornerstones of echocardiography and to this day are the most widely used imaging modes for diagnosis of a wide range of cardiovascular diseases. Artificial intelligence is now playing a critical role in streamlining quantification of these two imaging modes in Vivid scanners.

1. What is Artificial Intelligence and how can this improve clinical workflows?

Artificial intelligence is a loosely defined term that is used to refer to different types of algorithms (see Glossary of Terms at the end of this document for further details). In this white paper when talking about **Deep Learning**, we refer to the ability of a machine to **Self-Learn** without implicit instruction.

The beauty of this approach is that the machine learns from data, in the same way that a human learns from experience by pairing observations with labels (to learn languages for example). As such, the more data the machine "sees" the more it learns and thus the better it is able to perform.

Deep Artificial Neural Network algorithms have the potential to outperform previous machine learning algorithms, since machine learning algorithms require instruction from humans on which features in the data are important. In echocardiography, and indeed in medicine in general, some relationships or important features are still unknown and thus this group of self-learning methods has potential to go beyond what we know today. The data tells the story.

At the same time, understandability of these algorithms is critical and therefore transparency of their design and how they arrive to the given result is critical to understand, and eventually trust, the results. The description of the algorithms below is written accordingly to reach this goal.

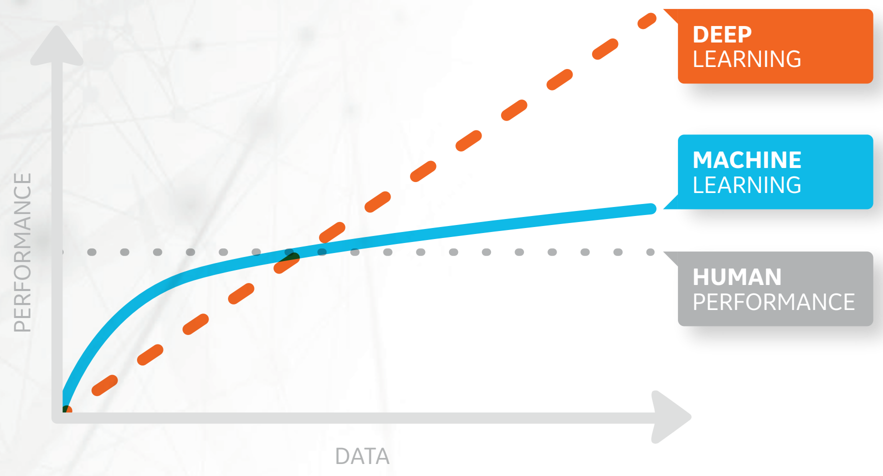


Fig. 1: Deep Learning has the potential to significantly improve accuracy and robustness over time as it "sees" more data.

2. Snapshot vs. Continuous Learning

Self-learning algorithms by construction get better with time. However, they require supervision to make sure that the learning goes in the right direction.

Vivid scanners are equipped with snapshot algorithms – this means that the performance of the algorithm is known and controlled. The self learning is done during the design process and then is set prior (snapshot) to release of the product. Thus the Vivid scanners are not continuously learning in the field. The risk with continuously learning algorithms is that they may start to perform irregularly if they receive data that can add confusion, for example due to different users performing the task differently.

Algorithms are updated with new data under controlled conditions, in the same sense that one wouldn't want their teenager to learn to drive from a formula one race car driver, the scanners should not learn from any "driver".

3. Self-Learning Best practice

Self-learning approaches are data driven. Careful data handling is paramount in the development of accurate and robust algorithms.

Training an algorithm to perform well in the environment the algorithm was trained on is relatively easy (e.g. test on data coming from the same hospital the algorithm was trained on). The real challenge is to develop algorithms that work well (i.e. they are accurate) in most clinical settings (i.e. they are robust).

Accurate and robust algorithms are tested by taking an algorithm trained on data from one or more hospitals and applying to data from other hospitals (preferably several).

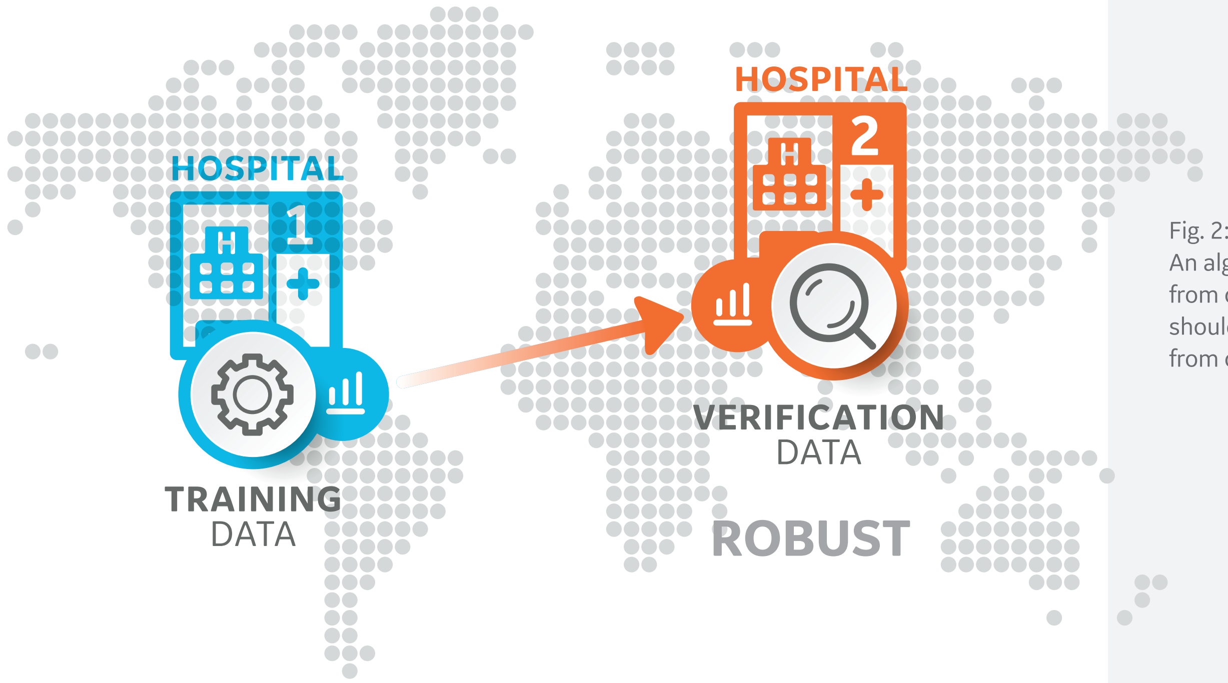


Fig. 2: An algorithm trained on data from one (or more) hospitals should work as well on data from other hospitals

Semi-automatic detection of the appropriate measurement

4. AI Auto Measure - Spectrum Recognition

Tools for semi-automatically measuring spectral Doppler images have been available on Vivid systems in recent years in the Cardiac Auto Doppler measurement package. Conventionally, the measurements were invoked by the user by entering the measurement menu and selecting the appropriate measurement given the image on screen. To circumvent the need to select the measurement, an AI algorithm was trained to semi-automatically detect the appropriate measurement, enabling the system to fast-forward the path from scanning to measurements.

Approach:

The approach chosen for this task was designed to mirror the way a human would perform this same task – which essentially boils down to the following: given an existing Doppler spectrum image, how can the intended Doppler spectrum measurement be deduced? A human would typically look at the 2D image to see which valve or wall the Doppler cursor was positioned over. They would then use that information in conjunction with the imaging mode, continuous wave (CW), pulsed wave (PW), or tissue velocity Doppler (TVD), and the shift of the baseline (indicating whether the user was interested in the positive part of the spectrum or the negative part). The imaging mode and baseline shift are parameters stored in every Doppler file (and thus do not need to be predicted), the remaining task is to detect which valve or wall the Doppler cursor was located on by the user.

The approach was therefore to train an AI algorithm on 2D images to semi-automatically detect where the Doppler Spectrum is being recorded (which valve, vessel, or ventricular wall). A direct approach was taken to solve this task by feeding the 2D image to a Deep Learning classification network. In addition to the image layer, an additional image layer representing the Doppler gate location was included in the input to the network. Several state-of-the-art network architectures were tested, to find the optimal fitting configuration which gave the best accuracy.

This algorithm was tested on a verification dataset of thousands of images from a range of different hospitals around the world, different from the hospitals used to train the algorithm, giving 98% accuracy and 100% reproducibility. [10]

The new workflow combining AI Auto Measure – Spectrum Recognition with the Doppler spectrum measurements (manual or automated with Cardiac Auto Doppler) is shown in Fig. 3.

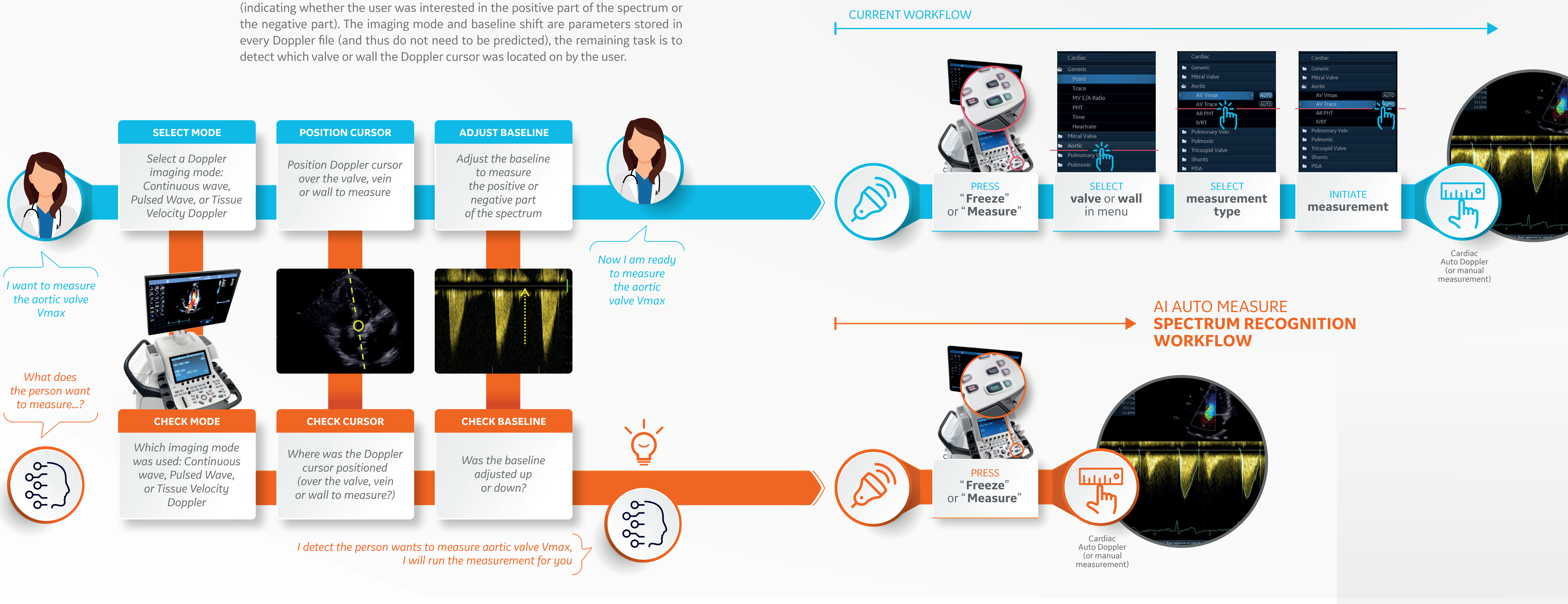
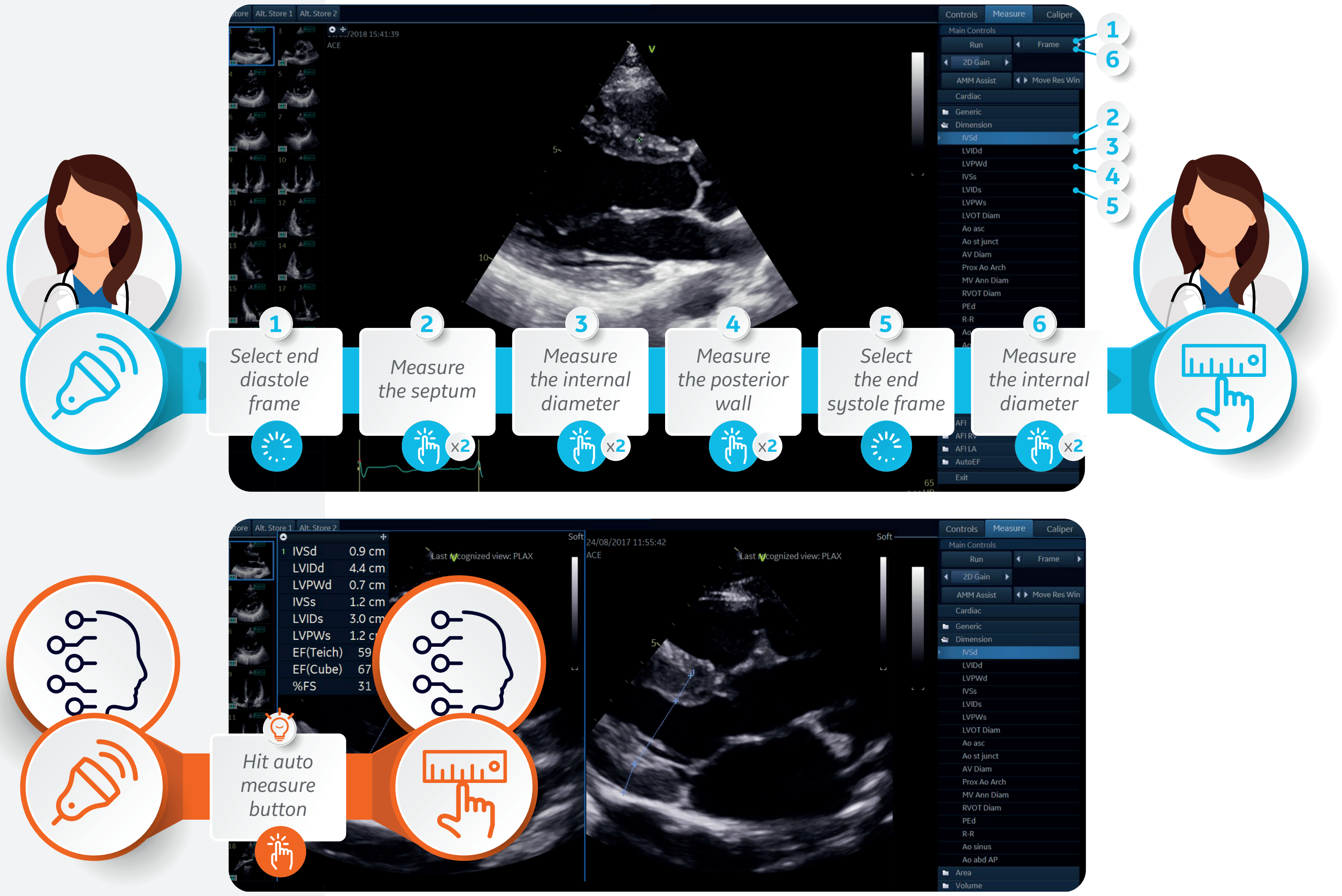


Fig. 3: The workflow impact from the AI Auto Measure – Spectrum Recognition algorithm, reducing the need to choose the appropriate valve or wall to measure and then subsequently invoke the relevant measurement for that valve or wall (top row, manual workflow), to streamline the process.

5. AI Auto Measure – 2D

A known limitation of 2D echocardiography measurements is user variability. Measurements of the left ventricle are time-consuming and known to be challenging due to difficulties in distinguishing the structures of the right heart from the septum and in distinguishing the posterior wall from papillary muscles in the left ventricle. Moreover, the guidelines on how to perform these measurements, are not always consistently followed in daily practice. Thus, the system is embedding a deep neural network (NN) which can detect the relevant points in the image, from which the following measurements of the left ventricle are derived:

- Intra-ventricular septal thickness (IVS)
- Left-ventricular internal diameter (LVID)
- Left-ventricle posterior wall thickness (LVPW)



Approach:

The approach chosen to address this task was to train a Deep Artificial Neural Network algorithm to predict the end points of each measurement (i.e. a landmark detection approach). An x and y channel were appended to the 2D one-channel (grayscale) image to enable a coordinate convolution layer to be applied, which provides additional spatial context to a convolution neural network (see Fig. 5). Several state-of-the-art network architectures were tested, with the final chosen architecture being the one that gave the best accuracy. The algorithm was tested on hundreds of images and performed comparable to, or better than, human users with 100% reproducibility.

6. AI-based View Recognition

The goal of AI-based View Recognition is to automatically detect which standard 2D scan plane is acquired and store this label in the image file to be used later for streamlining workflows (for example, detecting if an image is suitable for a given measurement as described in the previous section or for automatically selecting a trio of compatible apical images for strain analysis as described further in the next section). The algorithm is now able to recognise most standard scan planes.

Approach:

A direct approach was taken to solve this task by feeding 2D image loops from standard view planes with corresponding image labels to a Deep Artificial Neural Network. Several state-of-the-art network architectures were tested, with the final chosen architecture being the one that gave the best accuracy. The input to the network was a one-channel grayscale image.

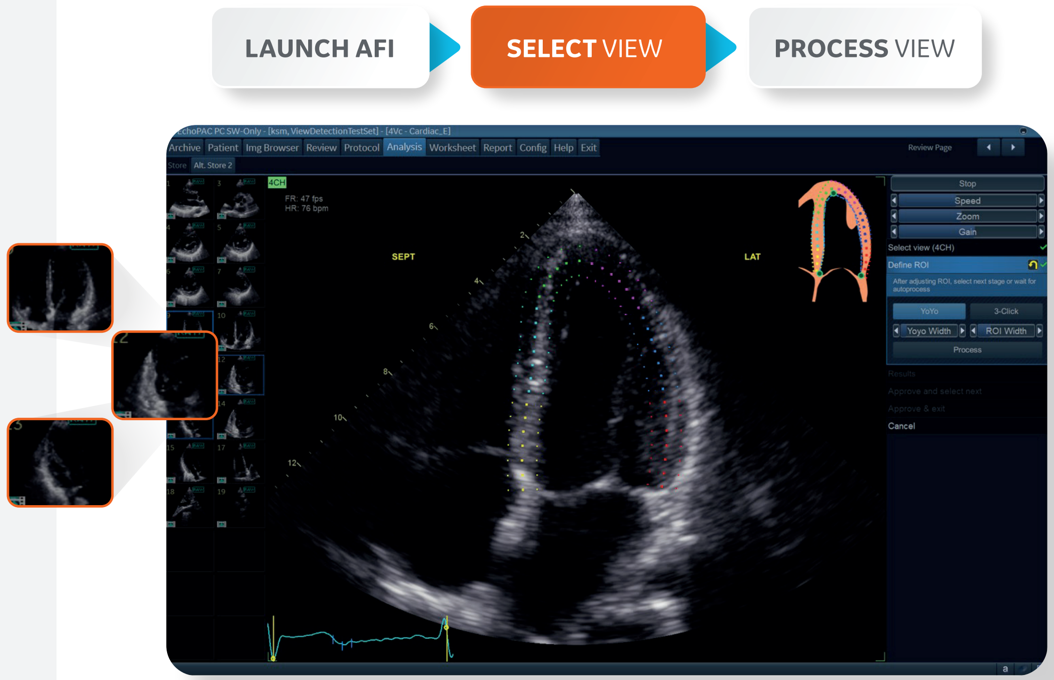
The most frequently acquired views from the parasternal, apical, and subcostal imaging windows were labelled for the algorithm. In some cases where subclassification of views was not necessary for existing functionality on the scanner, these were grouped to one label to improve performance. The final view labels cover the parasternal, apical, and subcostal imaging windows.

The algorithm was tested on a verification dataset of thousands of images from a range of different hospitals around the world, completely independent of the hospitals used to train the algorithm, giving 99% accuracy and 100% reproducibility.[9]

7. AI-enabled AFI LV

AI-based View Recognition, as described in the previous section, is now improving the workflow of AFI LV. Full strain analysis of the LV requires three apical views (the 4-chamberapical long axis, and 2-chamber views). These images must be compatible in terms of heart rate and frame rate in order for the strain calculations in each view to be able to be combined for a complete LV analysis. Matching strain calculations for views that vary significantly in terms of heart rate or frame rate will not be physiologically meaningful.

With AI-based View Recognition detecting the apical views the algorithm can combine the view information with the heart rate and frame rate, to automatically select a trio of apical images suitable for AFI LV analysis.



AI-enabled AFI LV: images have been pre-selected and labelled ready for processing

Conclusion

Ultra Edition AI algorithms help reduce operator fatigue

AI Auto Measure 2D helps reduce need for manual interactions (Up to 80% less clicks [11]), helping save up to 85% time on LV caliper measurements in EchoLab.[12]

Consistent measurements, reduced inter-operator dependency

With up to **99% accuracy** and **98% detectability** [9]**,** Ultra Edition AI algorithms provide consistent and repeatable measurements regardless of operator experience.

Ultra Edition Artificial Intelligence that assists you

Ultra Edition AI algorithms make the machine an integral part of the care team. Designed to relieve clinicians and technologists of the tedious yet complex task of various measurements; they help them direct their attention to the procedure, leveraging every second to focus their clinical expertise on what matters most, their patients; without compromising their own health.

UP TO
80%
LESS CLICKS [11]

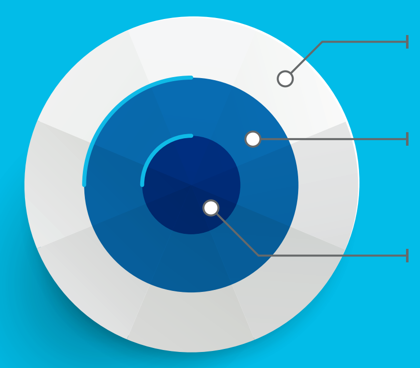
UP TO
85%
TIME SAVED [12]

99%
ACCURACY [9]

98%
DETECTABILITY [9]

100%
REPRODUCIBILITY [8,9,10]

Glossary



ARTIFICIAL INTELLIGENCE / SELF-LEARNING
Machines programmed to learn from data

MACHINE LEARNING
Machines capable of learning from defined features in data

DEEP LEARNING
Deep Artificial Neural Network algorithms capable of learning directly from data without explicit instruction on which features to focus on

Artificial Intelligence: The umbrella term that refers to the general notion of simulating the human learning process whereby machines are programmed to process data similar to human learning processes.

Self-Learning: The ability to learn without direct instruction.

Supervised/Unsupervised/Semi-supervised Learning: A programmed algorithm to learn common relationships between groups within a dataset by pre-grouping the data (supervised learning) or by automatically detecting groups within the data based on the data itself (unsupervised learning) or a combination of the two (learn from a few pre-grouped datapoints and extend to further ungrouped datapoints).

Reinforcement Learning: A programmed algorithm to learn optimal actions to lead towards a final goal by learning from feedback of each intermittent action.

Machine Learning: An algorithm to learn commonalities (and outliers) in data using supervised, unsupervised, semi-supervised, or reinforcement learning approaches.

Artificial Neural Network: An architecture of processing steps simulating brain neuron networks that combine several layers of data processing steps applied to data at varying levels of detail. An artificial neural network of more than two layers is called a Deep Neural Network.

Deep Artificial Neural Network: A subset of machine learning algorithms that automatically detect important features in the data (rather than being explicitly trained to focus on specific features). A set of algorithms that allow machines to seek relevant information itself in data without explicit direction thereby enabling significant improvements where machines can detect features in data that are unknown or unperceivable to a human user. The surge of Deep Learning applications is due to ongoing improvements in computer power that enable the use of increasingly deep Artificial Neural Network architectures which can potentially capture details in data not previously captured in other algorithms.

Algorithm: A pre-defined sequence of calculations.

Verification data: Data used to test the accuracy of an algorithm, collected from independent sites from those used to train an algorithm to ensure no overlap of patients and to ensure algorithms are not catered to specific sites, but rather that the algorithm extends well to other hospital settings.

Snapshot learning: An algorithm that is trained on data gathered at a fixed time point (a snapshot in time) under controlled conditions. The algorithm does not change on its own, and thus is 100% reproducible for a given software release.

Continuous learning: Infrastructure built-in to the device to continuously change the algorithm according to new data in an uncontrolled environment where there is potential for continuous improvement or degradation as the algorithm adapts to the new data it is fed with.

References:

1- https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

2- Forecasting the Future of Cardiovascular Disease in the United State, AHA Policy Statement, 2011, source: CIR.0b013e31820a55f5

3- Work Related Musculoskeletal Disorders In Sonography, SOCIETY OF DIAGNOSTIC MEDICAL SONOGRAPHY, Susan Murphey, https://journals.sagepub.com/doi/full/10.1177/8756479317726767

4- Temporal Trends in the Utilization of Echocardiography in Ontario, 2001 to 2009, Blecker et al, JACC: Cardiovascular Imaging Volume 6, Issue 4, April 2013, Pages 515-522, https://doi.org/10.1016/j.jcmg.2012.10.026

5- Focus cardiac ultrasound core curriculum and core syllabus of the European Association of Cardiovascular Imaging, Neskovic et al, European Heart Journal - Cardiovascular Imaging, Volume 19, Issue 5, May 2018, Pages 475–481 https://doi.org/10.1093/ehjci/jey006

6- Quality Measure and Quality Improvement, CMS.gov Centers for Medicare & Medicaid Services, https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/Quality-Measure-and-Quality-Improvement-

7- Transforming healthcare with AI. The impact on the workforce and organisations https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/transforming-healthcare-with-ai

8- Applicable to the AI Auto Measure - 2D algorithm, results based on GE internal data (DOC2367624).

9- Applicable to the AI-based View Recognition algorithm, results based on GE internal data (DOC2292732).

10- Applicable to AI Auto Measure - Spectrum Recognition, results based on GE internal data (DOC2292732).

11- Applicable to the AI Auto Measure - 2D algorithm, results based on GE internal data (DOC2361011).

12- Based on an internal GE study where three observers were doing both manual and automated caliper measurements of LV study parameters on 25 PLAX acquisitions. The achieved timing varied both for the automated and the manual measurements, but on average, over a total of 71 data points, the manual measurements took 76 seconds, while just 11 seconds with AI Auto Measure 2D

JB80498XX, JB000888XX