Automating Echocardiography Analysis using Deep Learning

Efficient measurement, workflow, and data generation

Andrew Gilbert

Disputas

02.07.2021

A global perspective of this thesis: *An automated analysis of structural echocardiography parameters using deep learning*

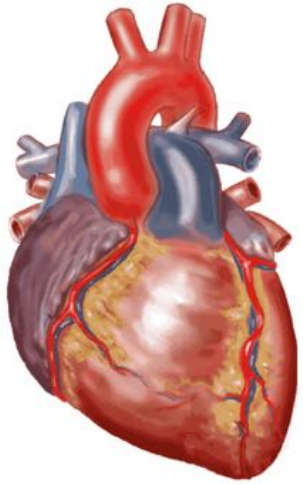
Goal: Make echocardiography more **efficient**, **reliable**, and **accurate**.

Clinical data

Mobile health monitor | Omics | Clinical reports | Clinical & experimental recordings | Medical images

Medical Imaging

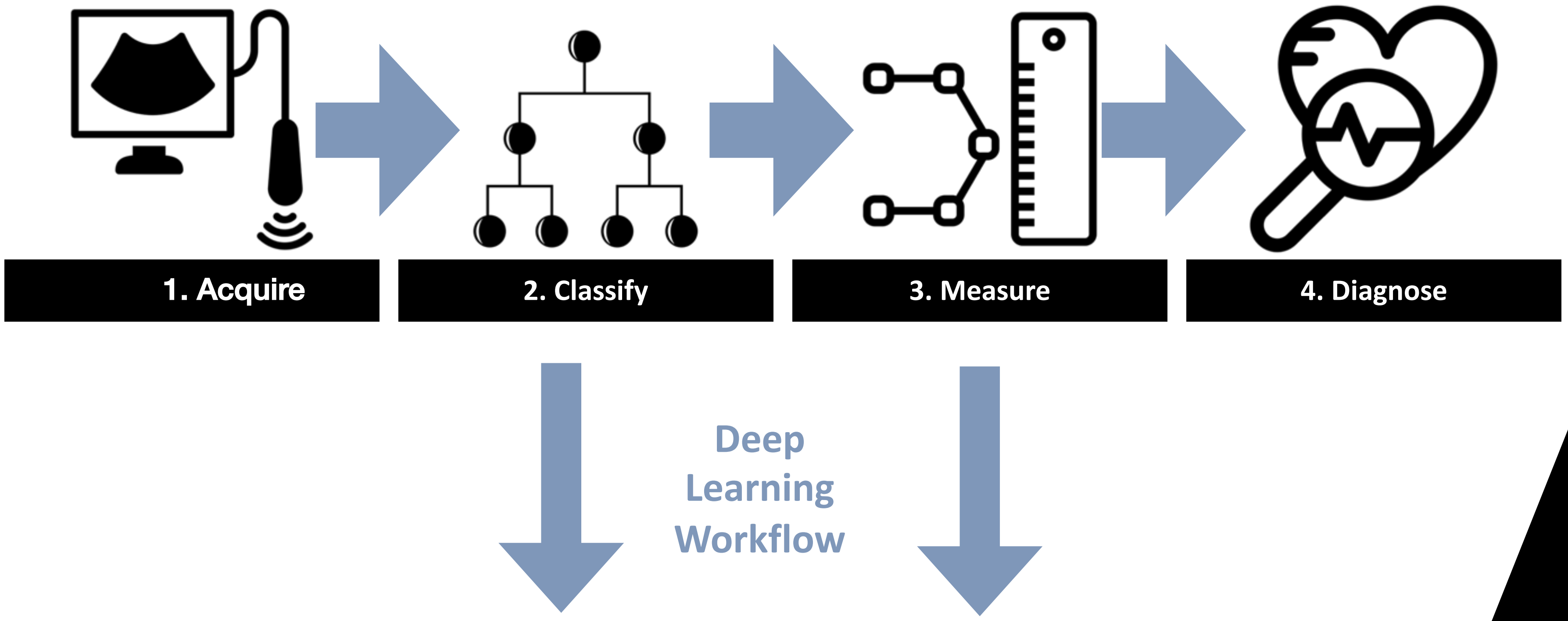
Nuclear imaging | Computed tomography | Magnetic resonance | Echo-cardiography



Echocardiography Analysis

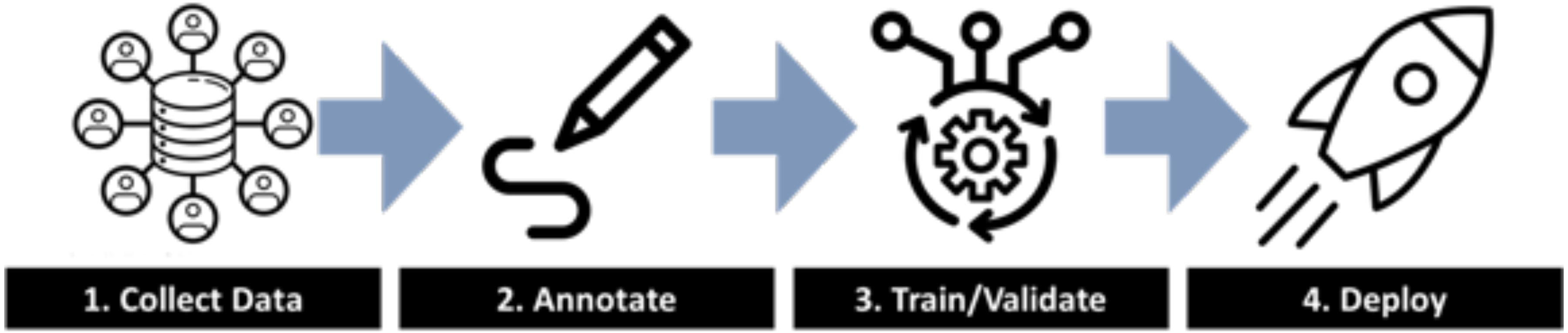
Qualitative analysis | Functional parameters | Structural parameters

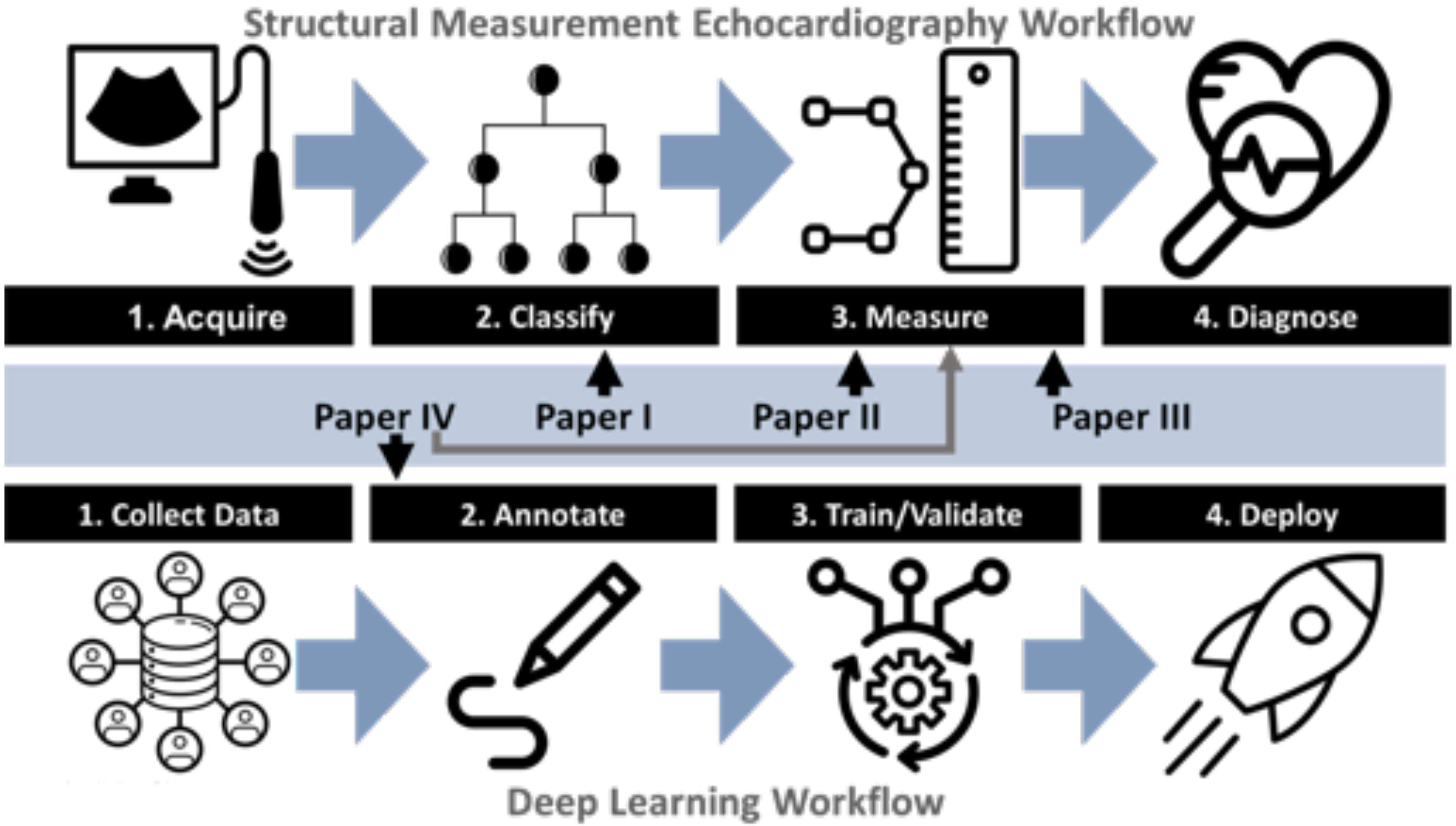
Structural Measurement Echocardiography Workflow



GE Healthcare
Universitetet i Oslo

Deep Learning Workflow





Structural Measurement Echocardiography Workflow
1. Acquire
2. Classify
3. Measure
4. Diagnose
Paper IV
Paper I
Paper II
Paper III
1. Collect Data
2. Annotate
3. Train/Validate
4. Deploy
Deep Learning Workflow

I: Doppler spectrum classification

II: Automated left ventricle dimension measurement

III: Curvature as a marker of basal septal hypertrophy

IV: Synthetic data generation

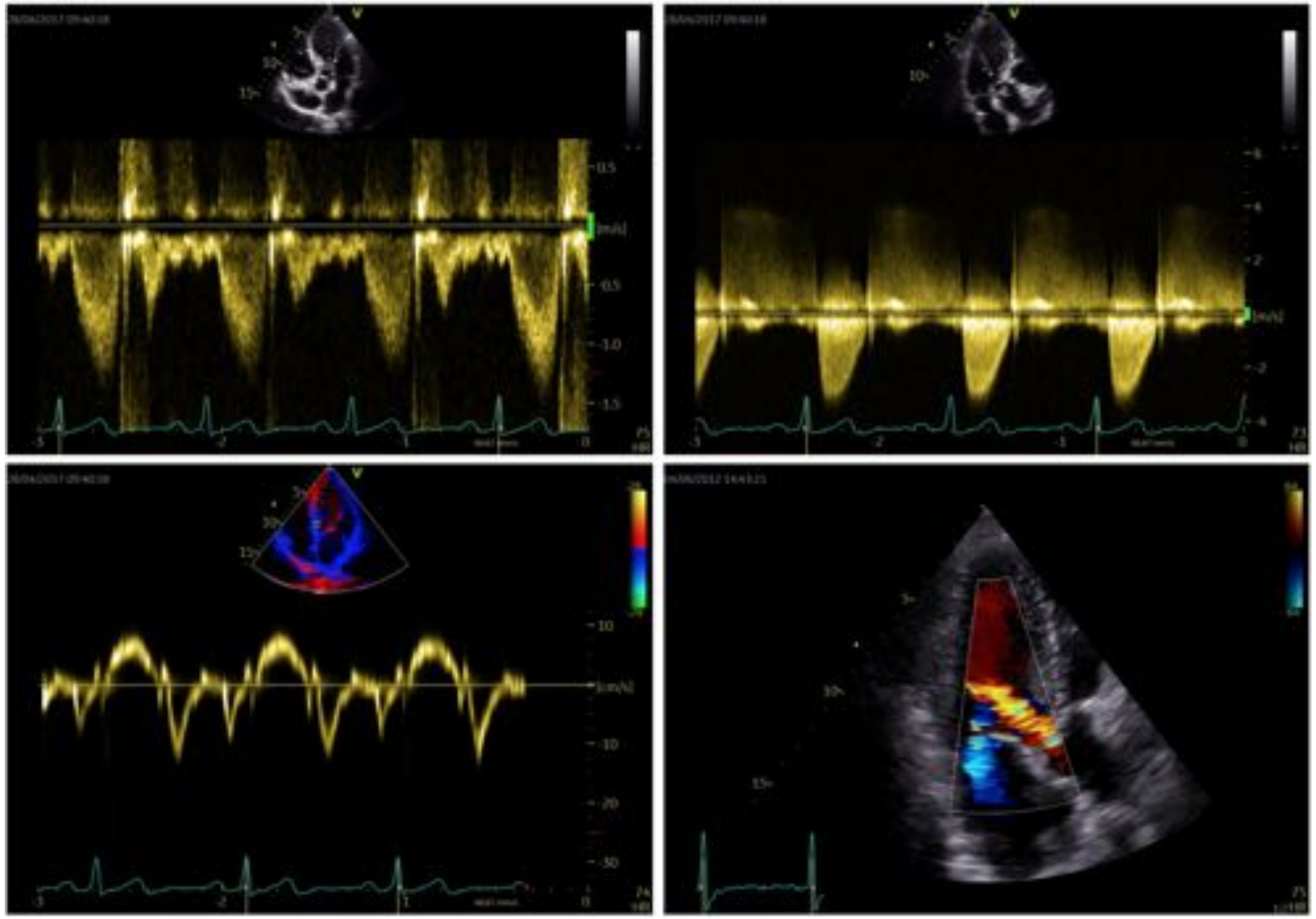
I. Doppler spectrum classification

Gilbert, A., Holden, M., Eikvil, L. Rakmail, M., Babić, A., Aaset, S. A., Samset, E., and McLeod, K. “User-Intended Doppler Measurement Type Prediction Combining CNNs with Smart Post-Processing”. In: *Journal of Biomedical and Healthcare Informatics*. (2020).

I: Doppler spectrum classification

Doppler imaging is an important modality for measurements of motion and flow.

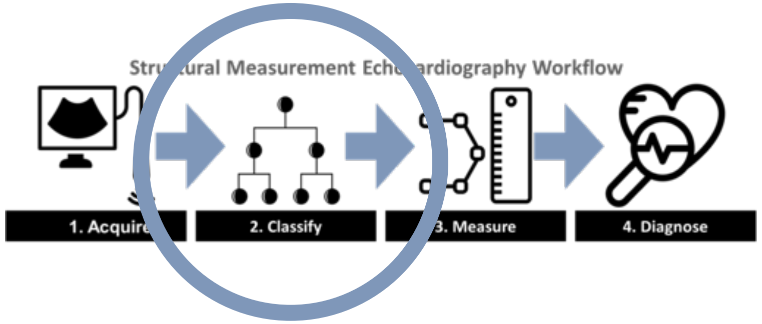
Pulsed Wave | Continuous Wave



Tissue Doppler | Color Flow

Doppler imaging can be used to measure blood flow and structural motion for many different regions

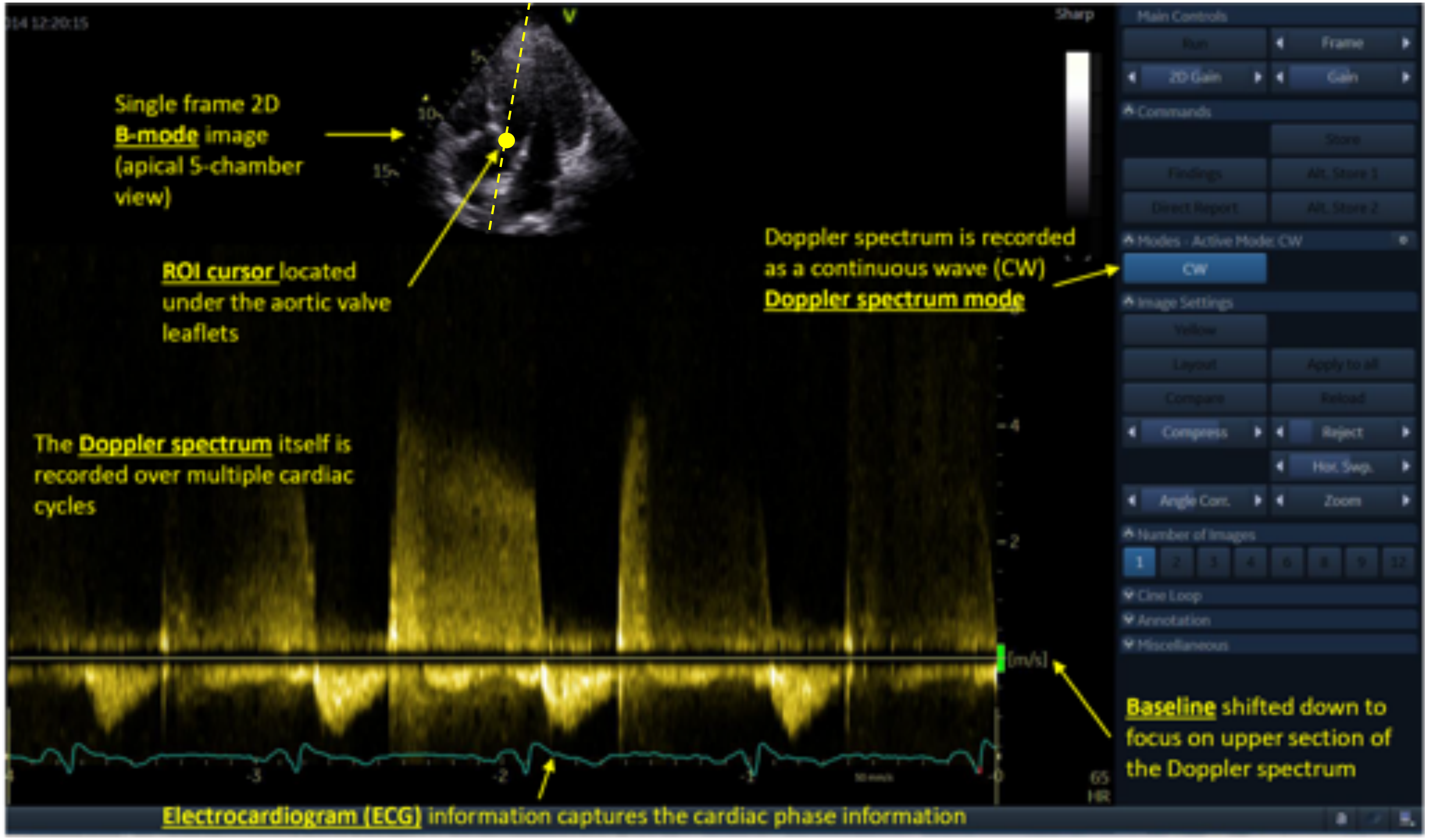
To measure, it is necessary to first identify which measurement should be performed



GE Healthcare
Universitetet i Oslo

There is a variety of information we can use to classify the image. Which are most relevant?

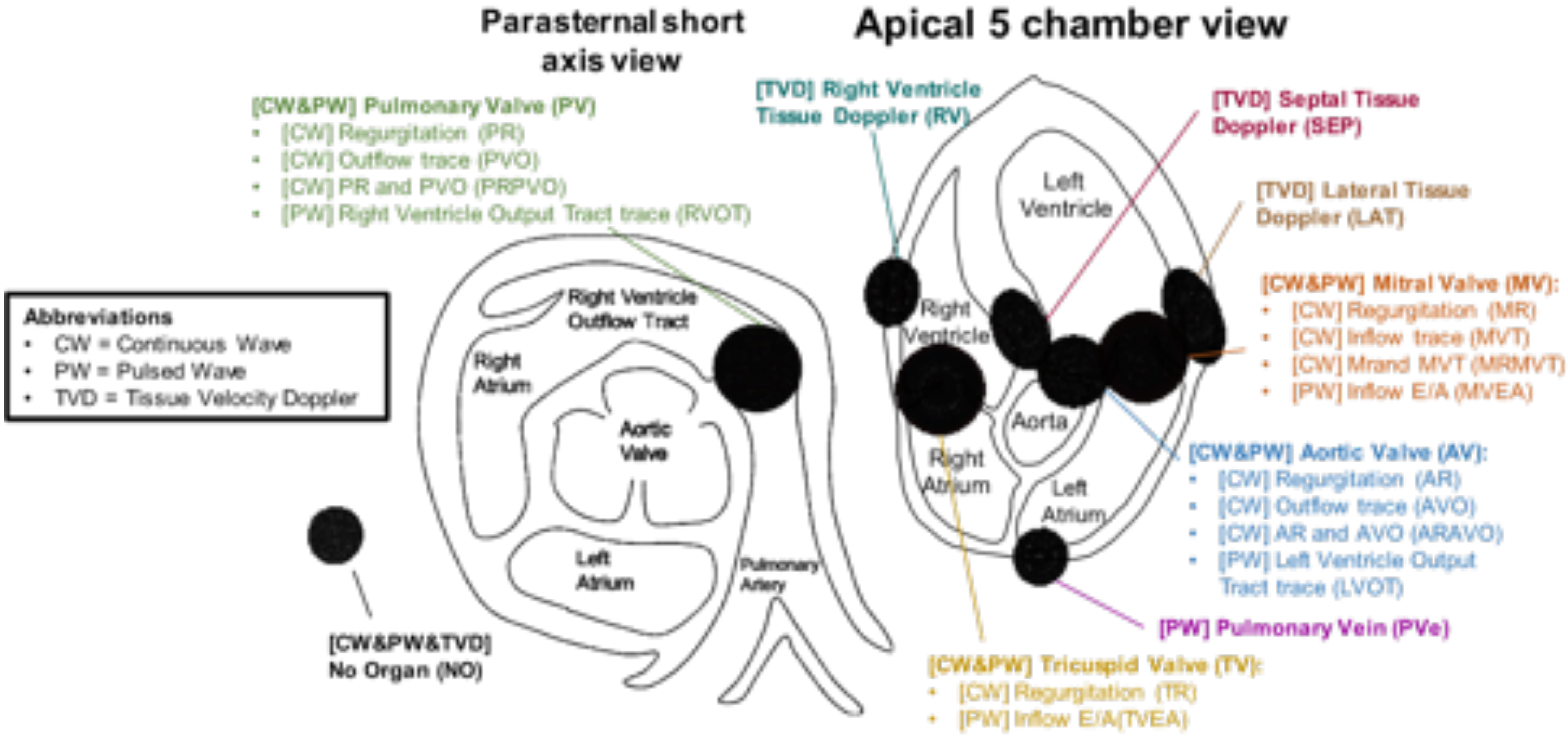
I: Doppler spectrum classification



GE Healthcare Universitetet i Oslo

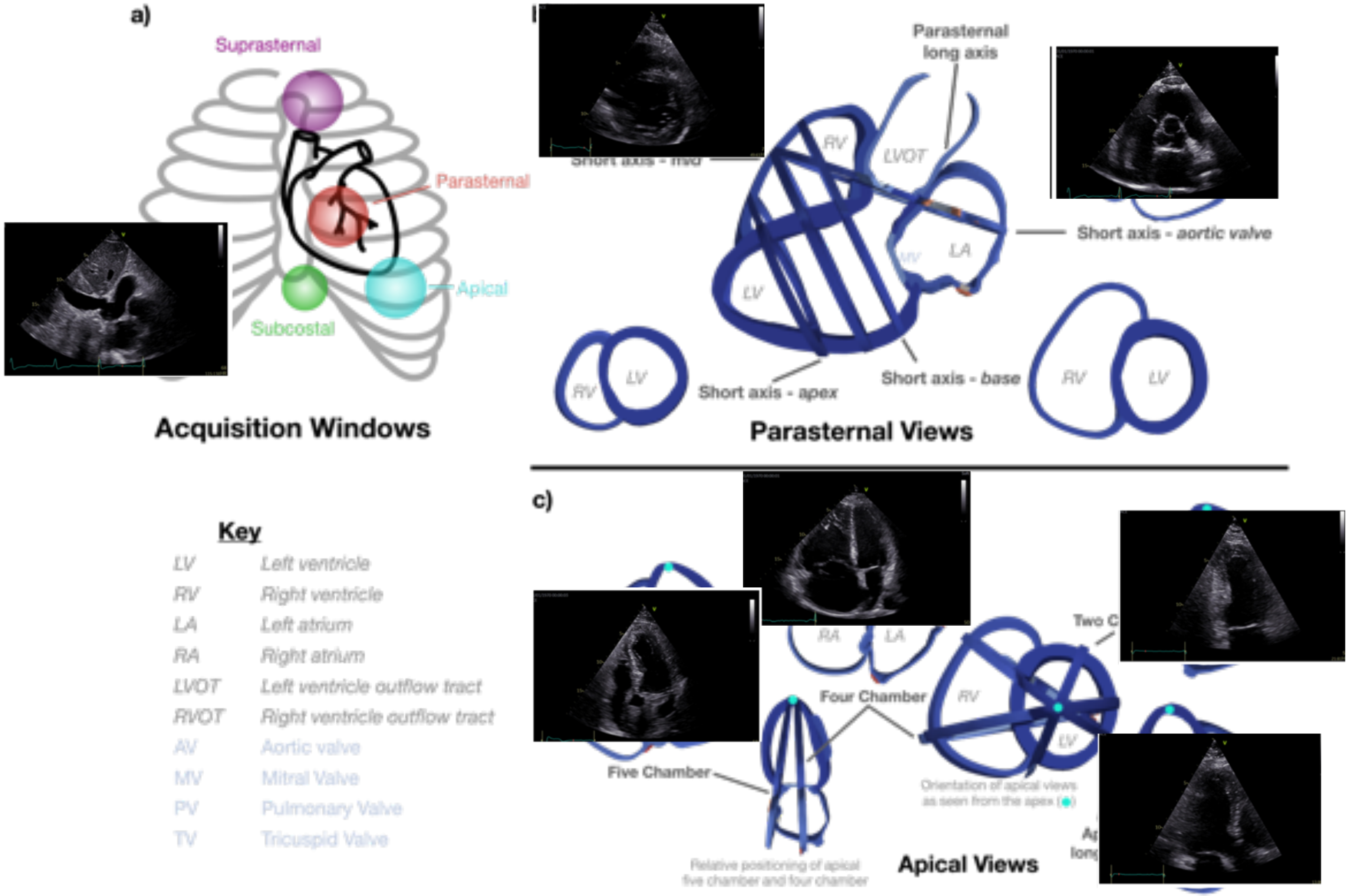
The measurement classes can be grouped by location

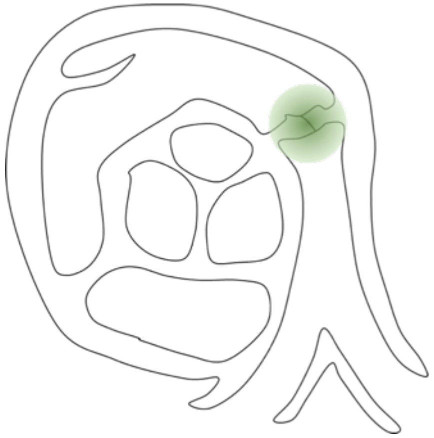
I: Doppler spectrum classification



Similar to view recognition, but B-mode and ROI information must be combined.

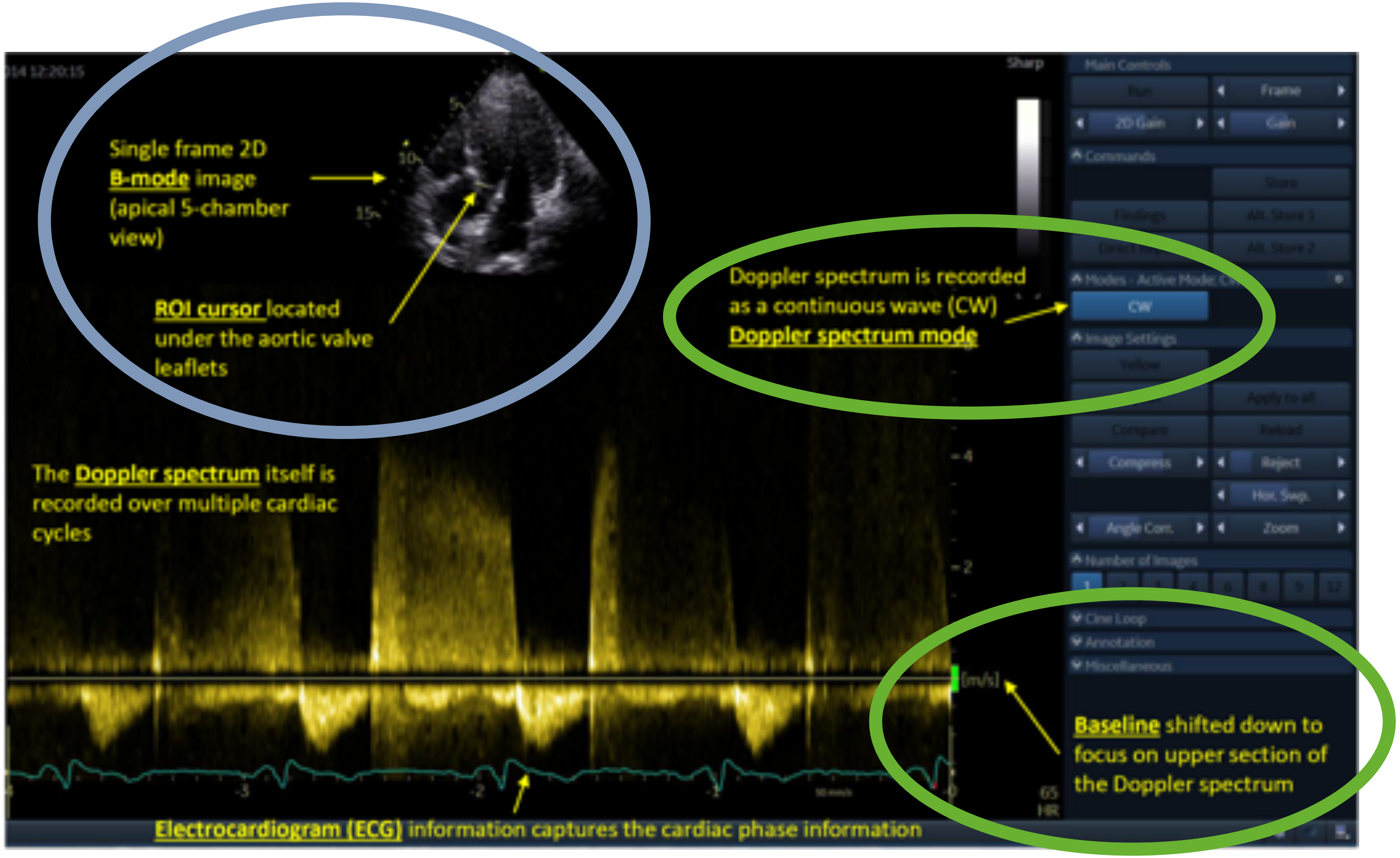
I: Doppler spectrum classification





There is a variety of information we can use to classify the image. Which are most relevant?

I: Doppler spectrum classification



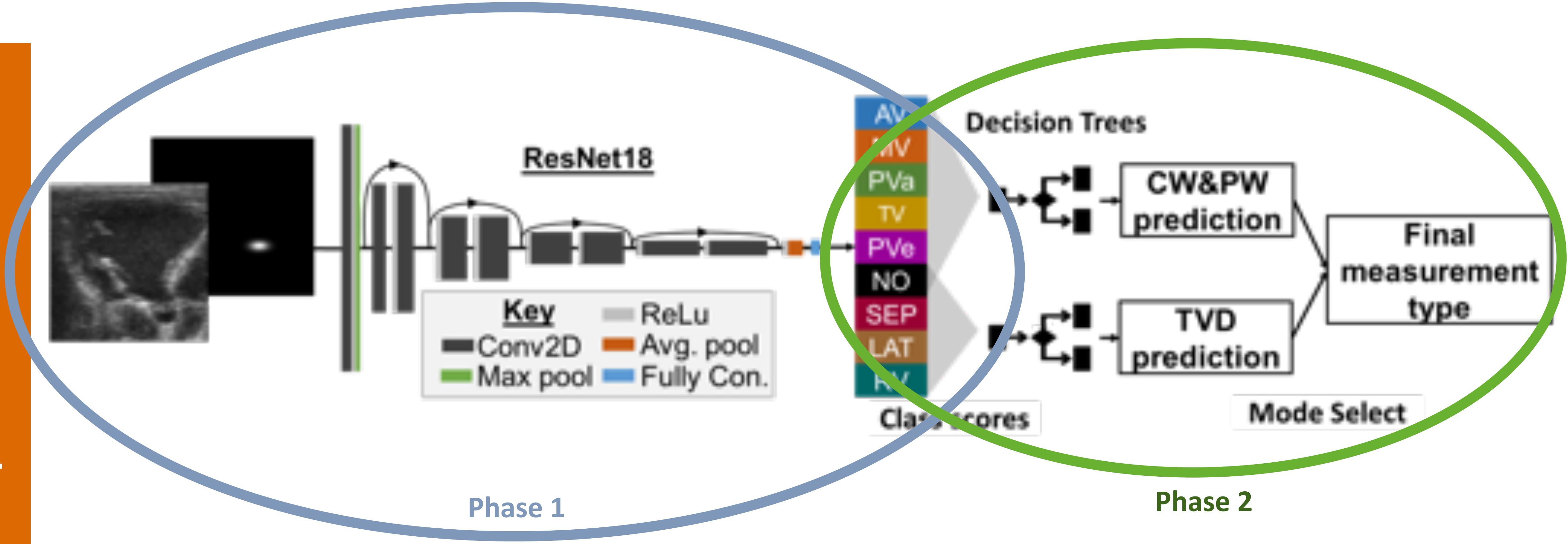
Phase 1: Identify the location of the measurement using the B-mode image and ROI cursor

Phase 2: Classify the measurement from the location (Phase 1) and other parameters

GE Healthcare Universitetet i Oslo

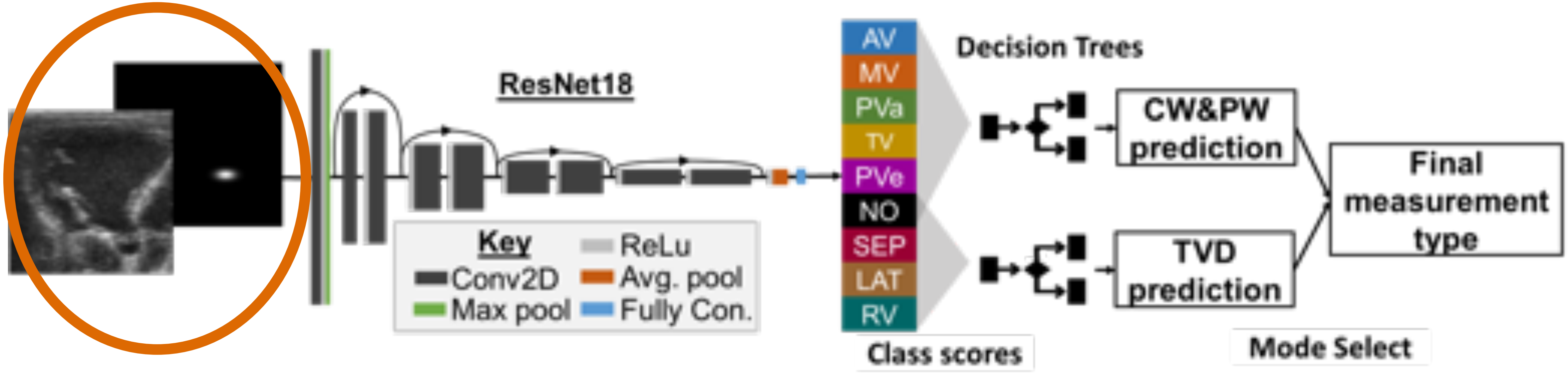
The network must learn to combine the imaging and ROI cursor information.

I: Doppler spectrum classification



The network must learn to combine the imaging and ROI cursor information.

I: Doppler spectrum classification

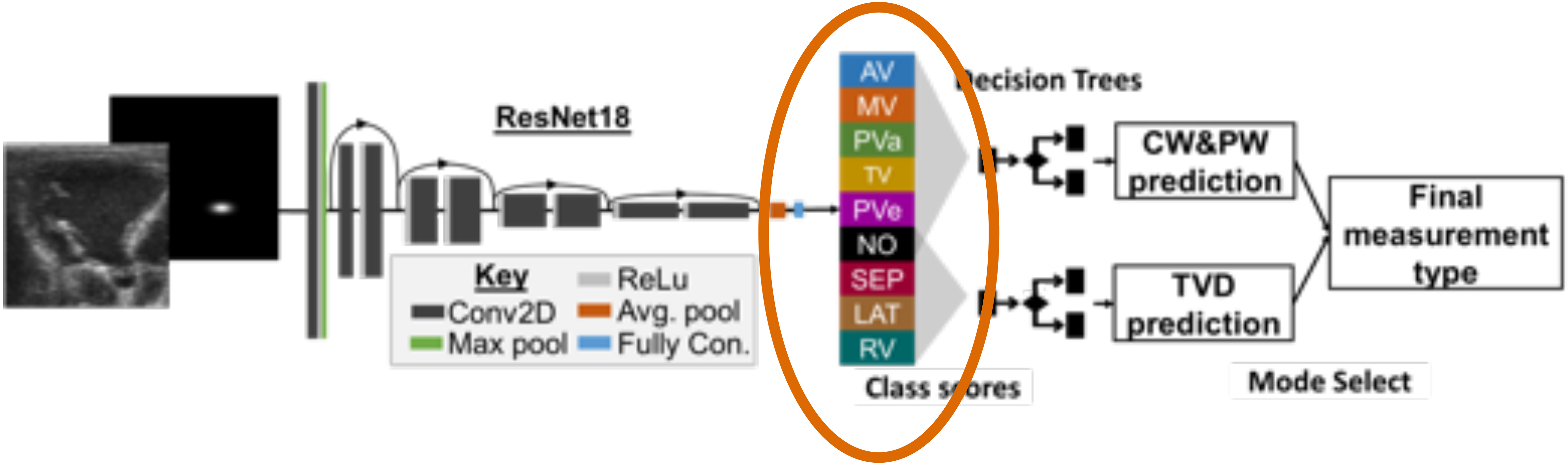


Key insights:

- **Heatmaps used to encode** location information at the input.
- **Multi-head learning** to integrate mode information.
- Thresholding and ensemble networks can be used to increase accuracy and as a **measure of confidence.**

The network must learn to combine the imaging and ROI cursor information.

I: Doppler spectrum classification

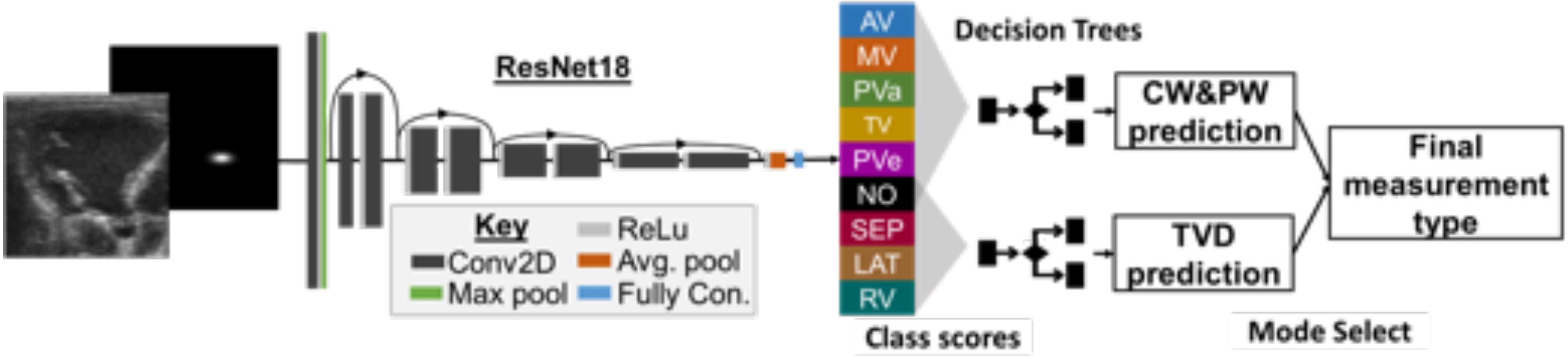


Key insights:

- **Heatmaps used to encode** location information at the input.
- **Multi-head learning** to integrate mode information.
- Thresholding and ensemble networks can be used to increase accuracy and as a **measure of confidence.**

The network must learn to combine the imaging and ROI cursor information.

I: Doppler spectrum classification

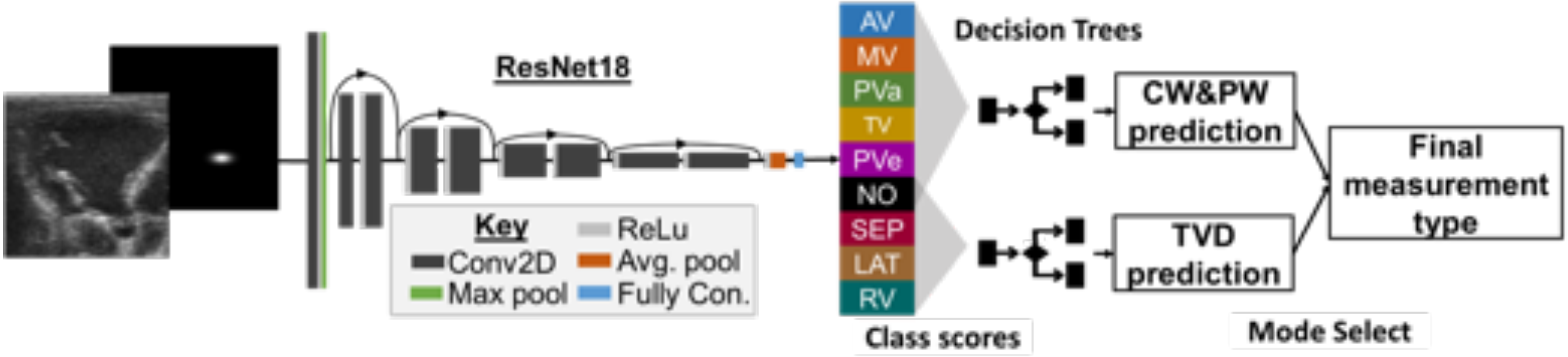


Key insights:

- **Heatmaps used to encode** location information at the input.
- **Multi-head learning** to integrate mode information.
- Thresholding and ensemble networks can be used to increase accuracy and as a **measure of confidence.**

The network must learn to combine the imaging and ROI cursor information.

I: Doppler spectrum classification



Key insights:

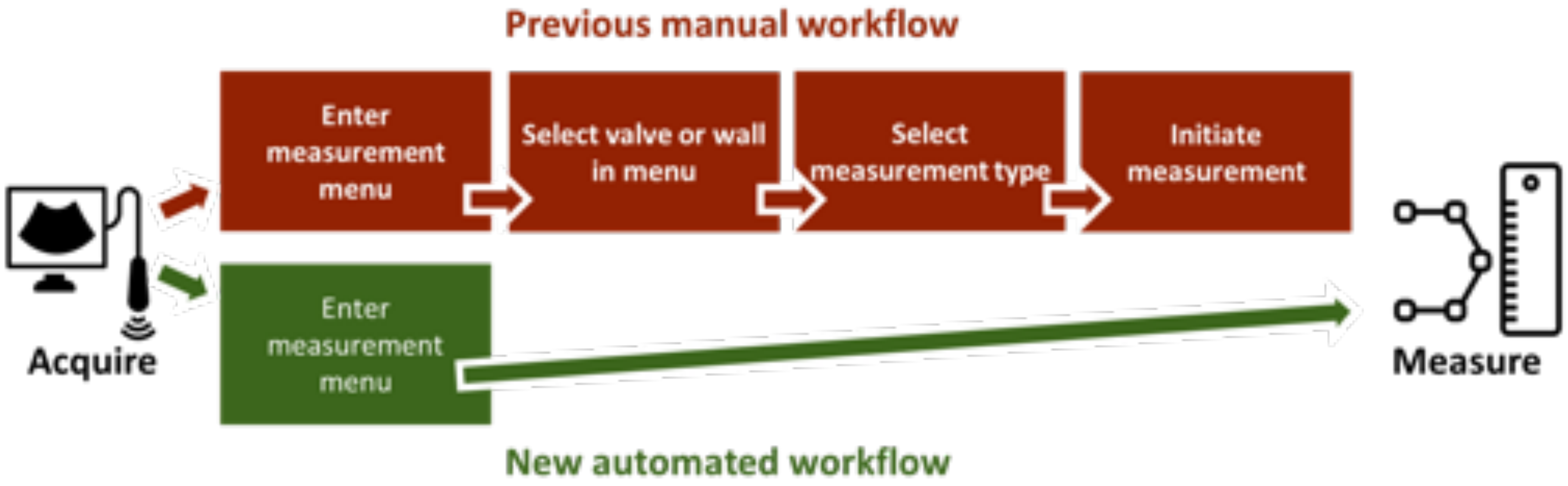
- **Heatmaps used to encode** location information at the input.
- **Multi-head learning** to integrate mode information.
- Thresholding and ensemble networks can be used to increase accuracy and as a **measure of confidence.**

Results:

- Achieve **96.4%** accuracy on a test set from a different distribution.
- Can be increased to **98.7%** accuracy using an ensemble of networks and ignoring 4.8% of "uncertain" images.

Automated classification brings workflow improvements on every Doppler scan

I: Doppler spectrum classification

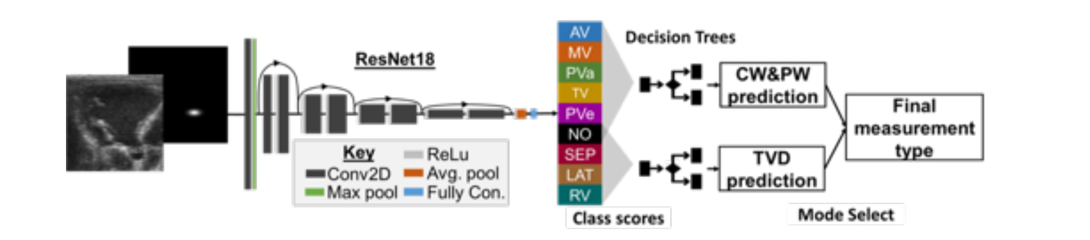


PIC
GE Healthcare
Universitetet i Oslo

Doppler spectrum classification implemented within Vivid Ultra Edition.

I: Doppler spectrum classification

I: Doppler spectrum classification



Highly accurate Doppler spectrum classification using **heatmap-encoding**, **multi-head training**, and **confidence metrics**.

II: Automated left ventricle dimension measurement

III: Curvature as a marker of basal septal hypertrophy

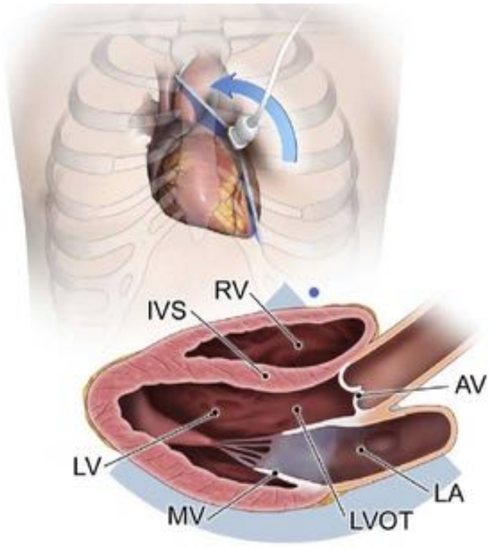
IV: Synthetic data generation

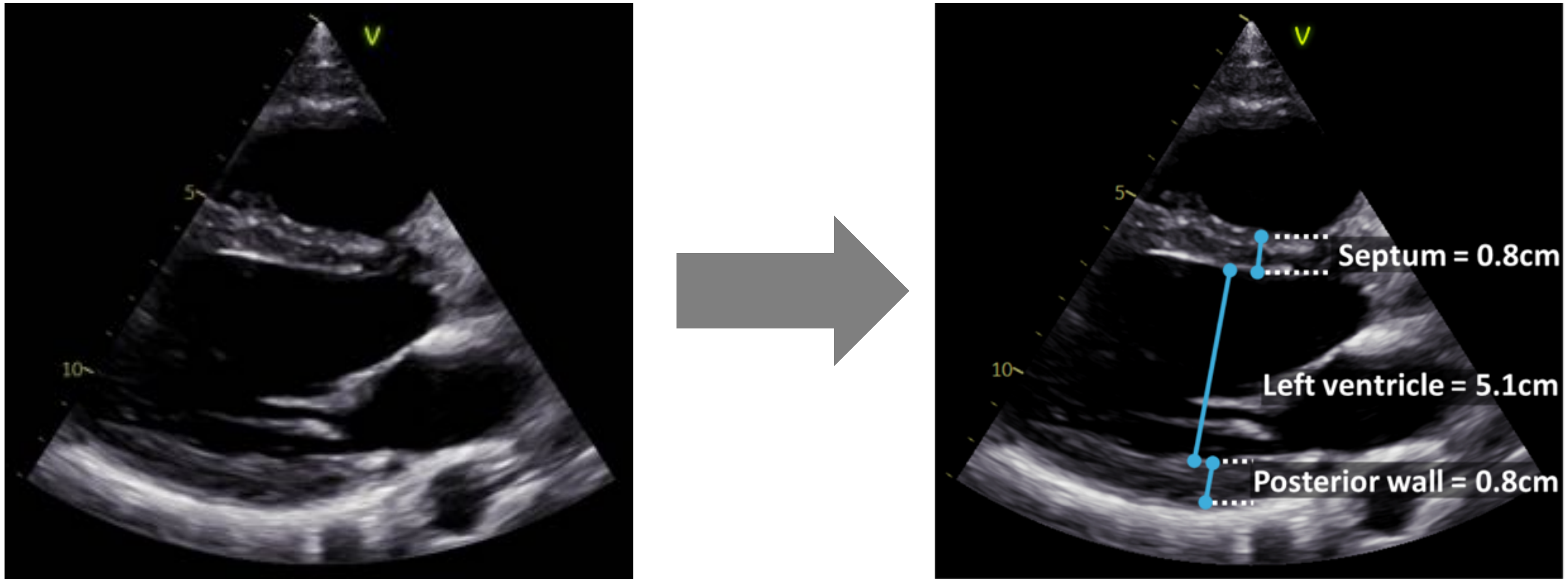
II: Left ventricle dimension measurement

Gilbert, A., Holden, M., Eikvil, L., Aase, S. A., Samset, E., and McLeod, K. “Automated Left Ventricle Dimension Measurement in 2D Cardiac ultrasound via an Anatomically Meaningful CNN Approach”. In: *Lecture Notes in Computer Science.* Vol. 11798, (2019), pp. 29-37.

Automate measurements of the left ventricle in the parasternal long axis (PLAX) view.

II: Left ventricle dimension measurement





Goal

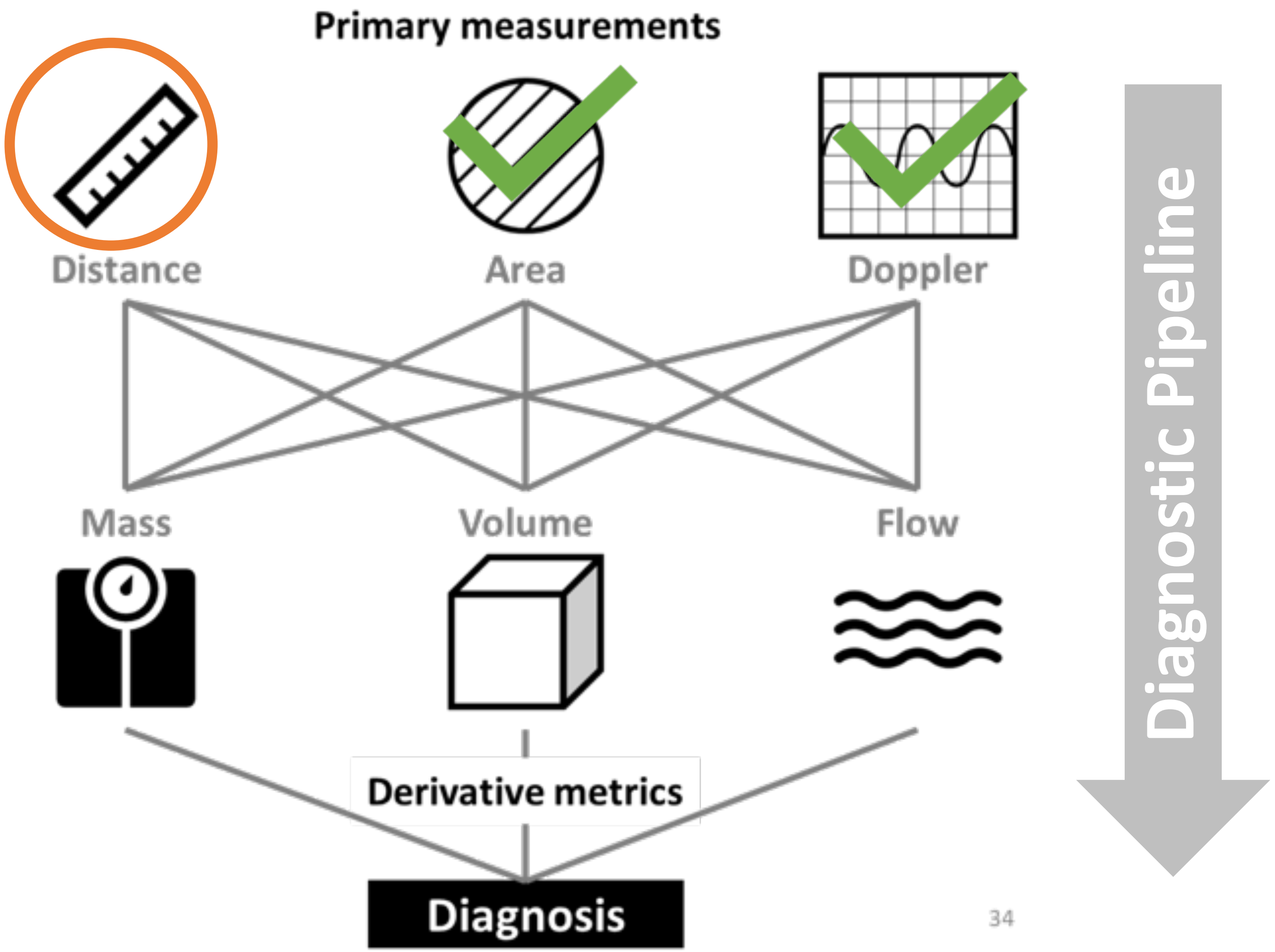
2D measurements of the left ventricle are a part of the first diagnostic measurements used for guiding patient treatments.

II: Left ventricle dimension measurement

High level of automation

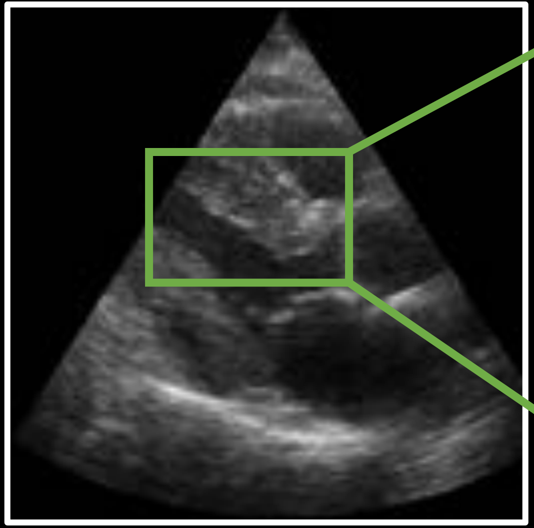
More automation needed

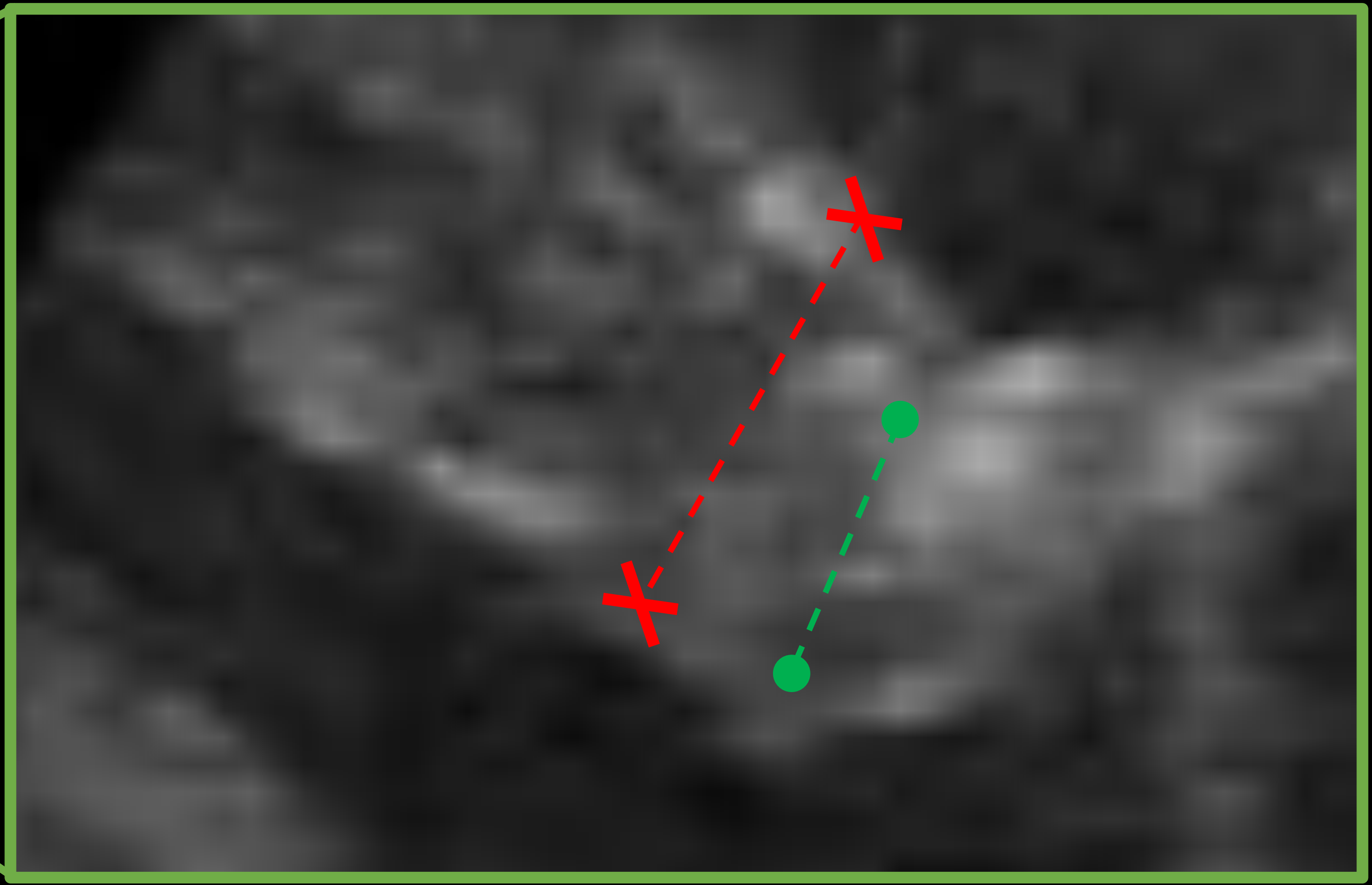
Echocardiography



2D measurements are often difficult to perform accurately

II: Left ventricle dimension measurement

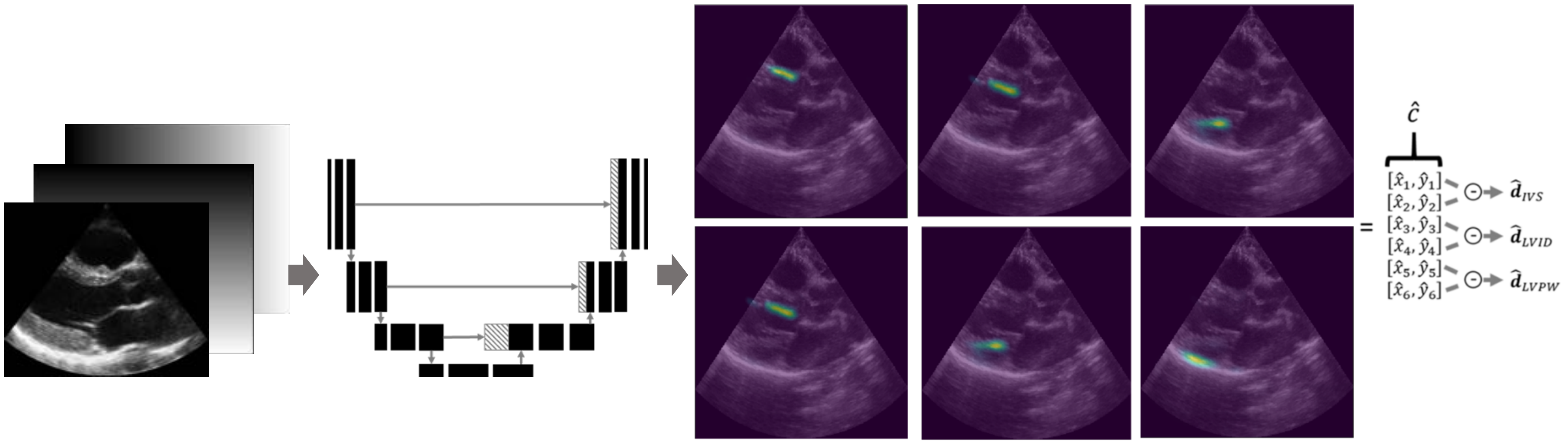




[cardioserv.net/6-pitfalls-accurate-lv-measurements]

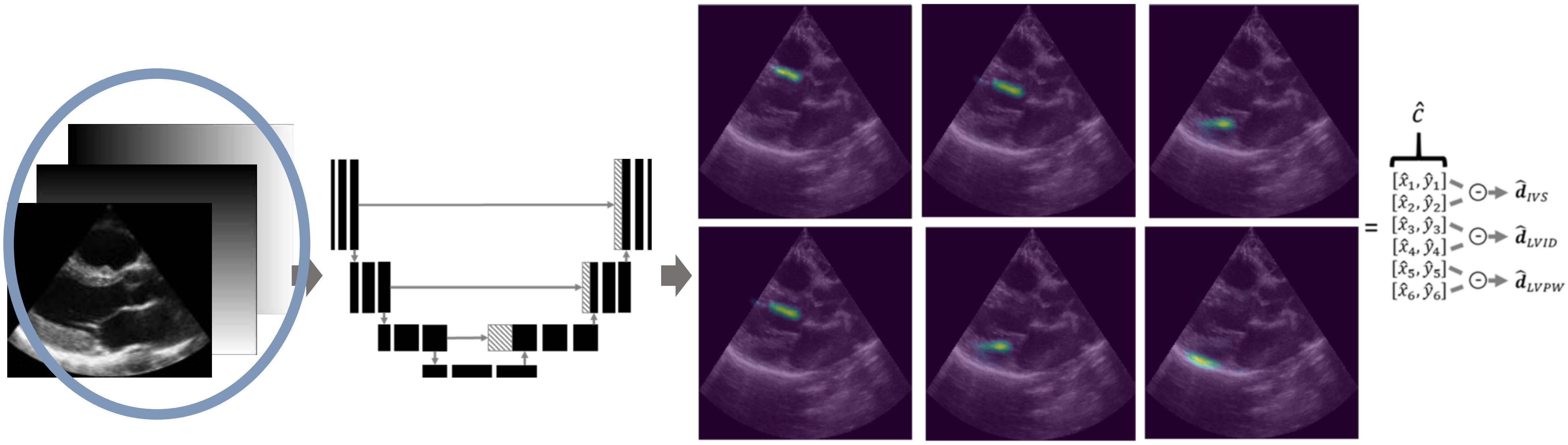
Framed as a landmark detection problem and solved with CNNs

II: Left ventricle dimension measurement



Framed as a landmark detection problem and solved with CNNs

II: Left ventricle dimension measurement

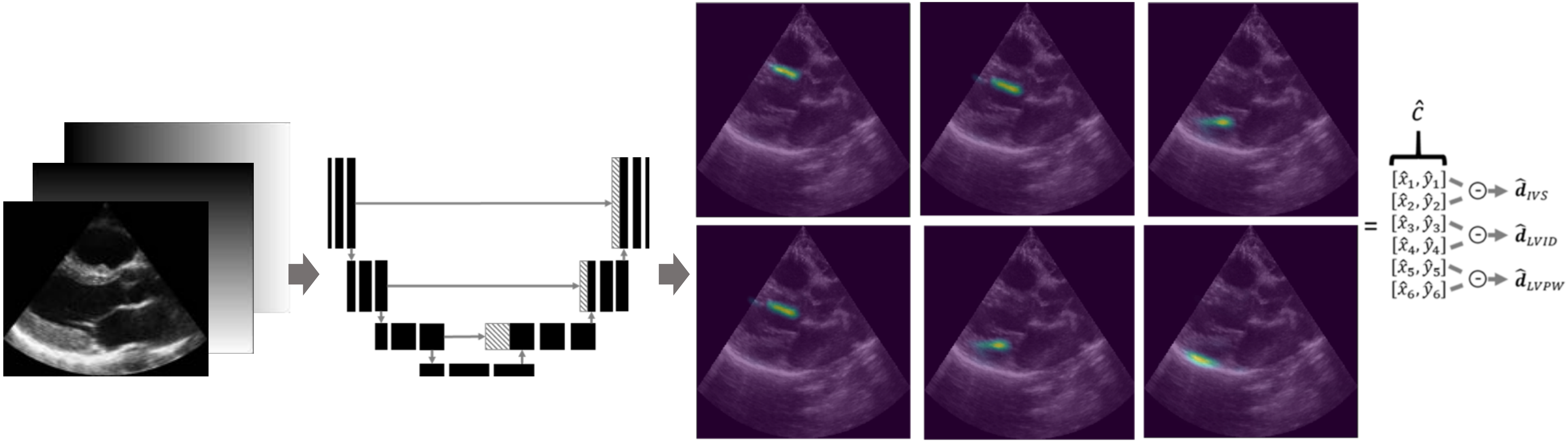


Key insights:

- **Coordinate convolution** can help distinguish between similar looking features.
- **Anatomical heatmaps** follow the expected distribution of the label uncertainty.
- A **multi-component loss function** helps optimize across multiple objectives

Framed as a landmark detection problem and solved with CNNs

II: Left ventricle dimension measurement



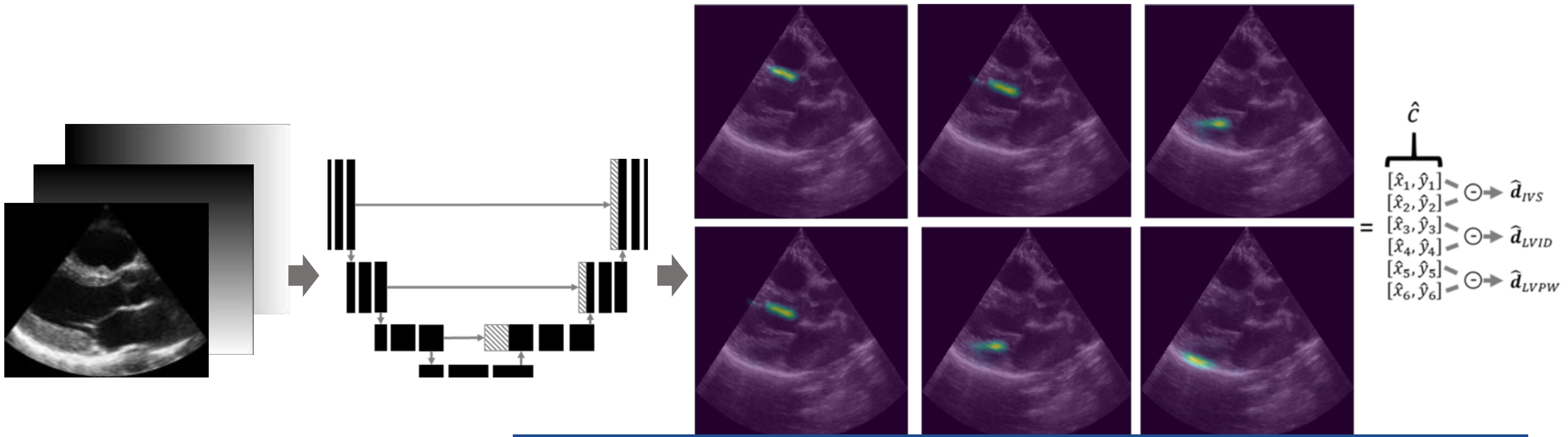
Key insights:

- **Coordinate convolution** can help distinguish between similar looking features.
- **Anatomical heatmaps** follow the expected distribution of the label uncertainty.
- A **multi-component loss function** helps optimize across multiple objectives

PIC GE Healthcare Universitetet i Oslo

Framed as a landmark detection problem and solved with CNNs

II: Left ventricle dimension measurement



Key insights:

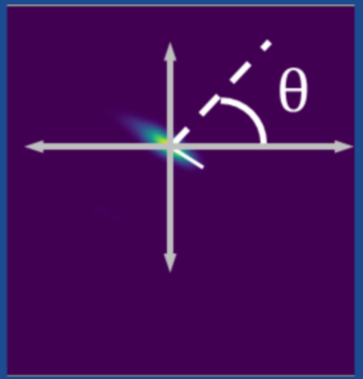
- **Coordinate convolution** (
- **Anatomical heatmaps** fo
- A **multi-component loss**

$$H_i \sim \mathcal{N}\left(\mu_i, \Sigma_j\right)$$

$$\mu_i = [x_i, y_i]$$

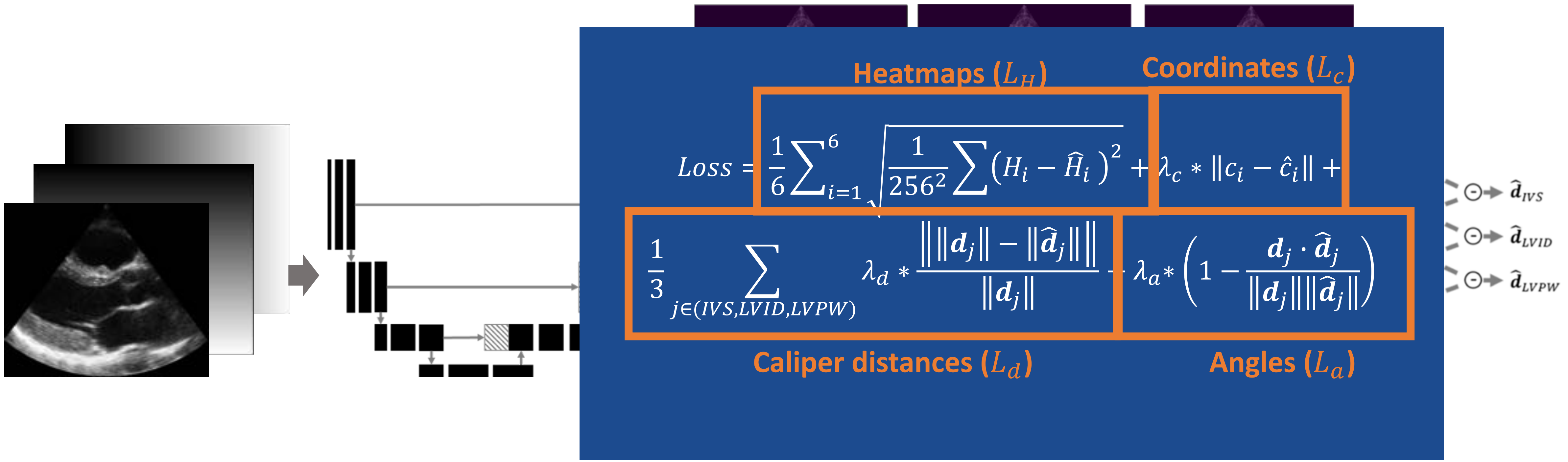
$$\Sigma_j = n * \left(R_j \cdot \begin{bmatrix} 1 & 0 \\ 0 & m \end{bmatrix} \cdot {R_j}^T\right)$$

R_j is a rotation matrix with $\theta_j = \angle \boldsymbol{d}_i$



Framed as a landmark detection problem and solved with CNNs

II: Left ventricle dimension measurement



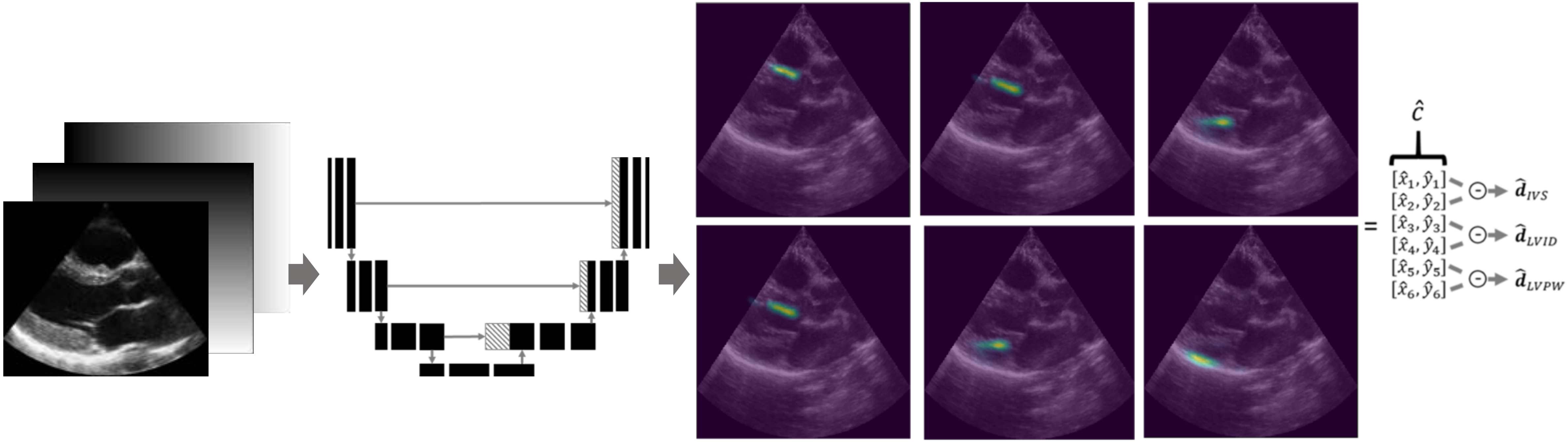
$$Loss = \frac{1}{6}\sum_{i=1}^{6}\sqrt{\frac{1}{256^2}\sum\left(H_i - \hat{H}_i\right)^2} + \lambda_c * \|c_i - \hat{c}_i\| + \frac{1}{3}\sum_{j\in(IVS,LVID,LVPW)} \lambda_d * \frac{\left|\|\boldsymbol{d}_j\| - \|\hat{\boldsymbol{d}}_j\|\right|}{\|\boldsymbol{d}_j\|} + \lambda_a * \left(1 - \frac{\boldsymbol{d}_j \cdot \hat{\boldsymbol{d}}_j}{\|\boldsymbol{d}_j\|\|\hat{\boldsymbol{d}}_j\|}\right)$$

Key insights:

- **Coordinate convolution** can help distinguish between similar looking features.
- **Anatomical heatmaps** follow the expected distribution of the label uncertainty.
- A **multi-component loss function** helps optimize across multiple objectives

Framed as a landmark detection problem and solved with CNNs

II: Left ventricle dimension measurement



Key insights:

- Coordinate convolution
- Anatomical heatmaps
- Multi-component loss function

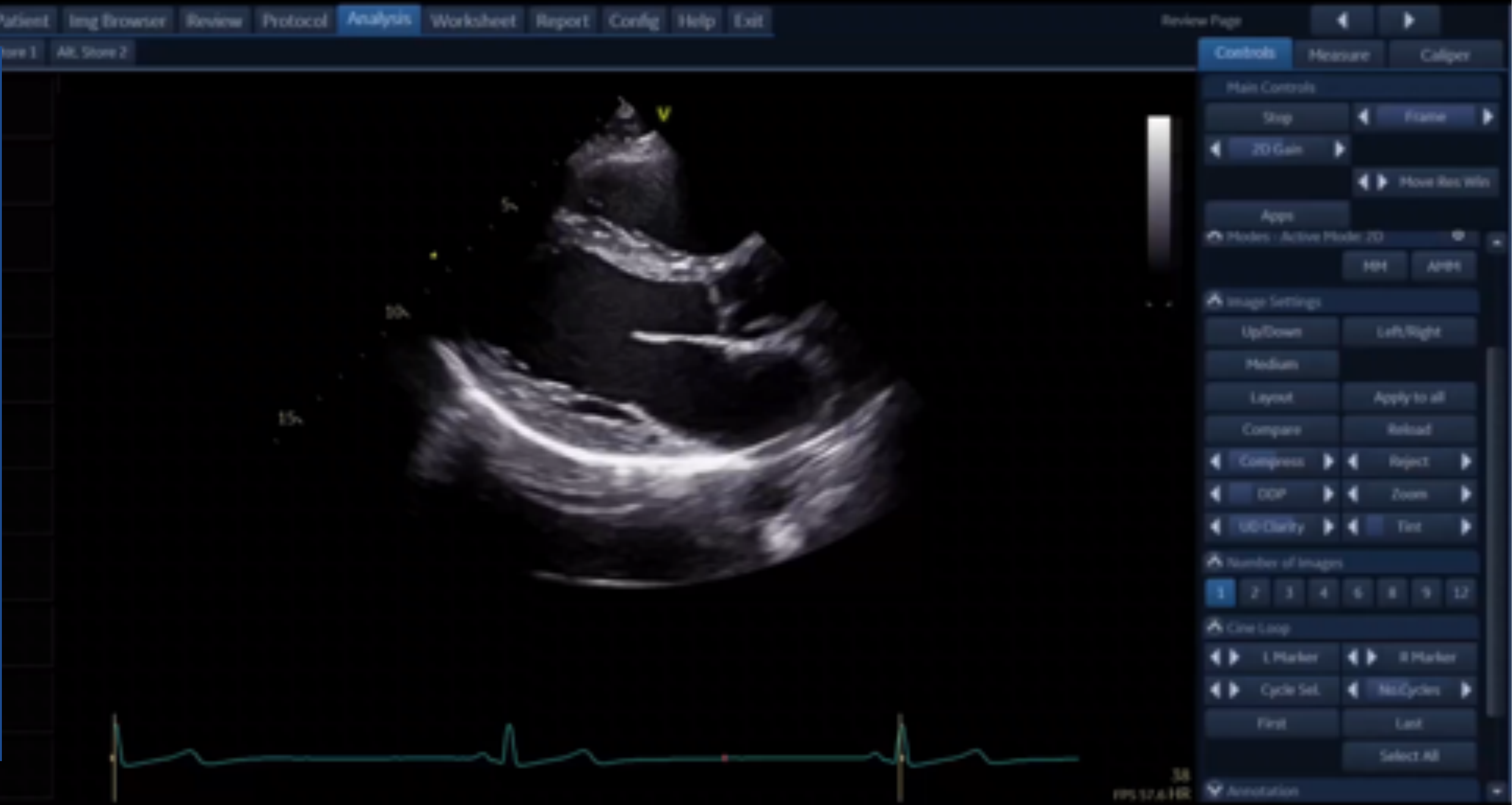
Accuracy approaches intra-analyzer error:

Method	Mean percent error
One expert annotating same image multiple times:	8.9%
Our approach:	10.0%

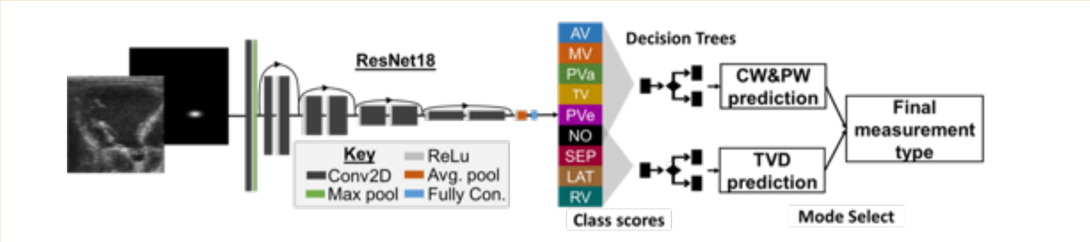
Left ventricle measurement also implemented within Vivid Ultra Edition.

II: Left ventricle dimension measurement

II: Left ventricle dimension measurement

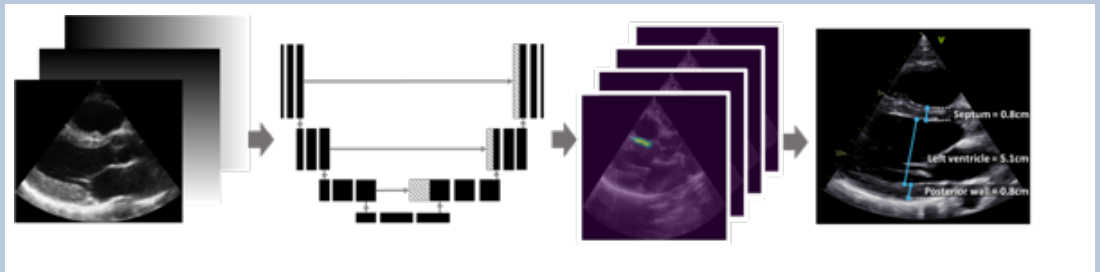


I: Doppler spectrum classification



Highly accurate Doppler spectrum classification using **heatmap-encoding**, **multi-head training**, and **confidence metrics**.

II: Automated left ventricle dimension measurement



Accurate two-dimensional measurements of the left ventricle using **coordinate convolution**, **anatomical heatmaps**, and a multi-component loss function.

III: Curvature as a marker of basal septal hypertrophy

IV: Synthetic data generation

III: Curvature as a marker of hypertrophy

Marciniak, M., **Gilbert, A.,** Loncaric, F., Fernandes, J. F., Bijnes, B., Sitges, M., King, A., Crispi, F., and Lamata, P. “Septal Curvature as a Robust and Reproducible Marker for Basal Septal Hypertrophy”. In: *Journal of Hypertension*. Vol. 38, (2021).

Basal septal hypertrophy affects 15-20% of hypertension patients.

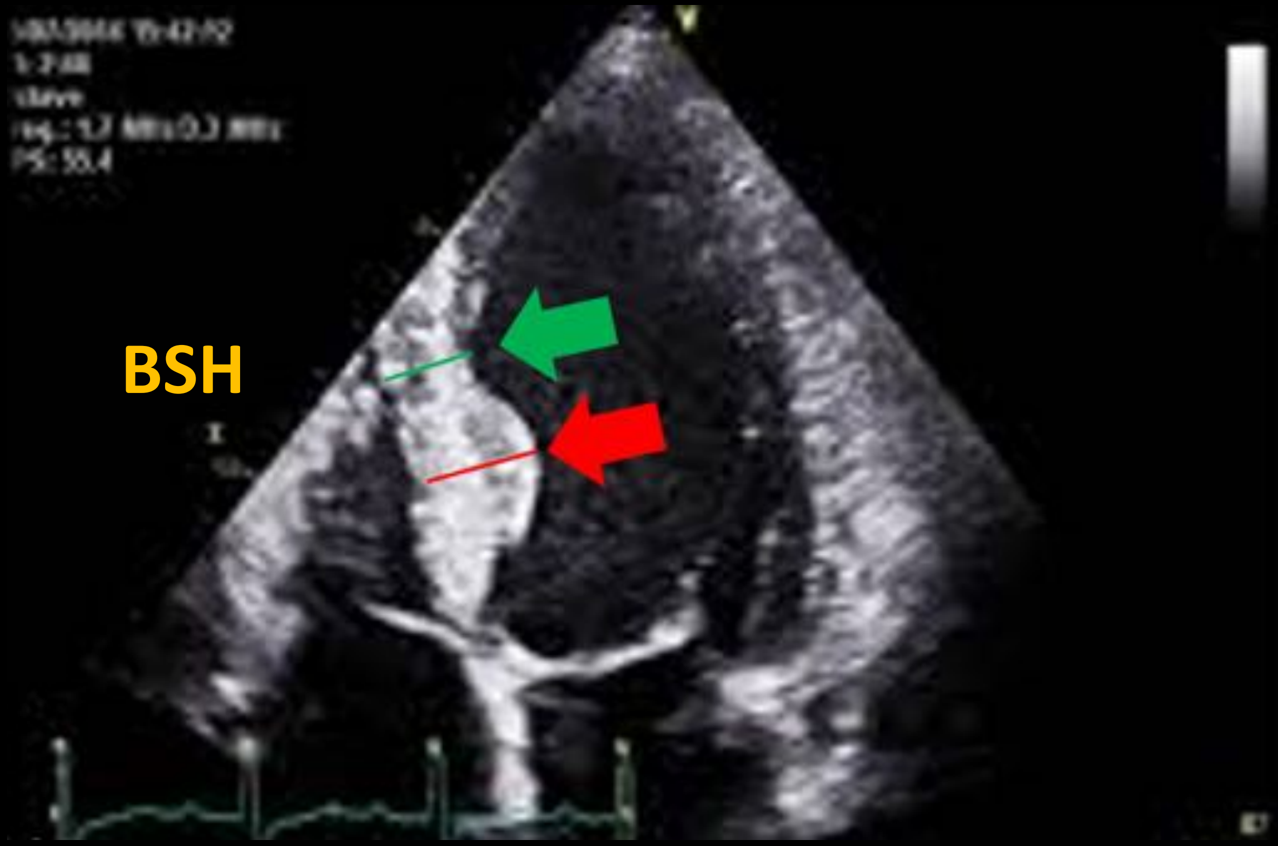
III: Curvature as a marker of hypertrophy

Basal septal hypertrophy (BSH) is an early marker of remodeling in hypertension.

Currently, diagnosed from wall thickness ratios and visual assessment.

Wall thickness ratios are difficult to reproduce.

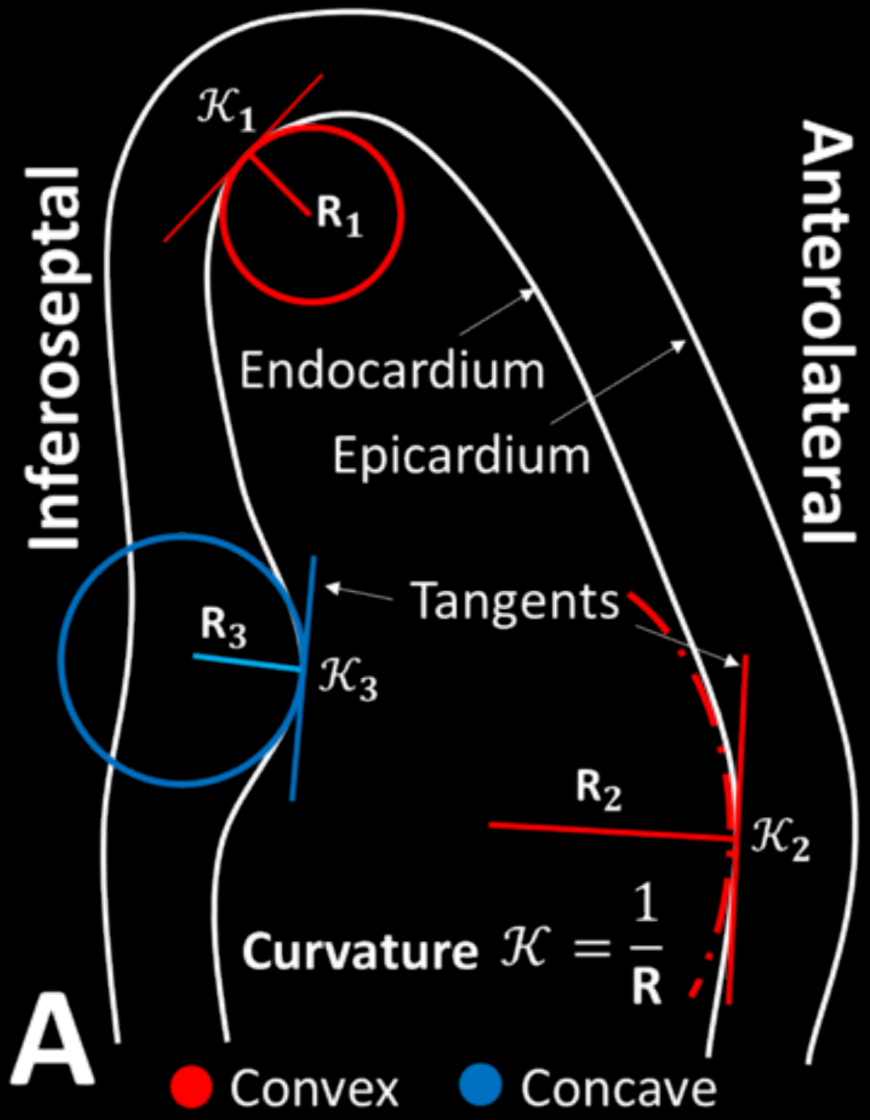
Visual assessment is reliable, but qualitative, we need a metric that quantifies visual assessment.



$$WTR = \frac{WT_{BASE}}{WT_{MID}}$$

The curvature of the endocardium can be quantified using the radii of tangent circles.

III: Curvature as a marker of hypertrophy

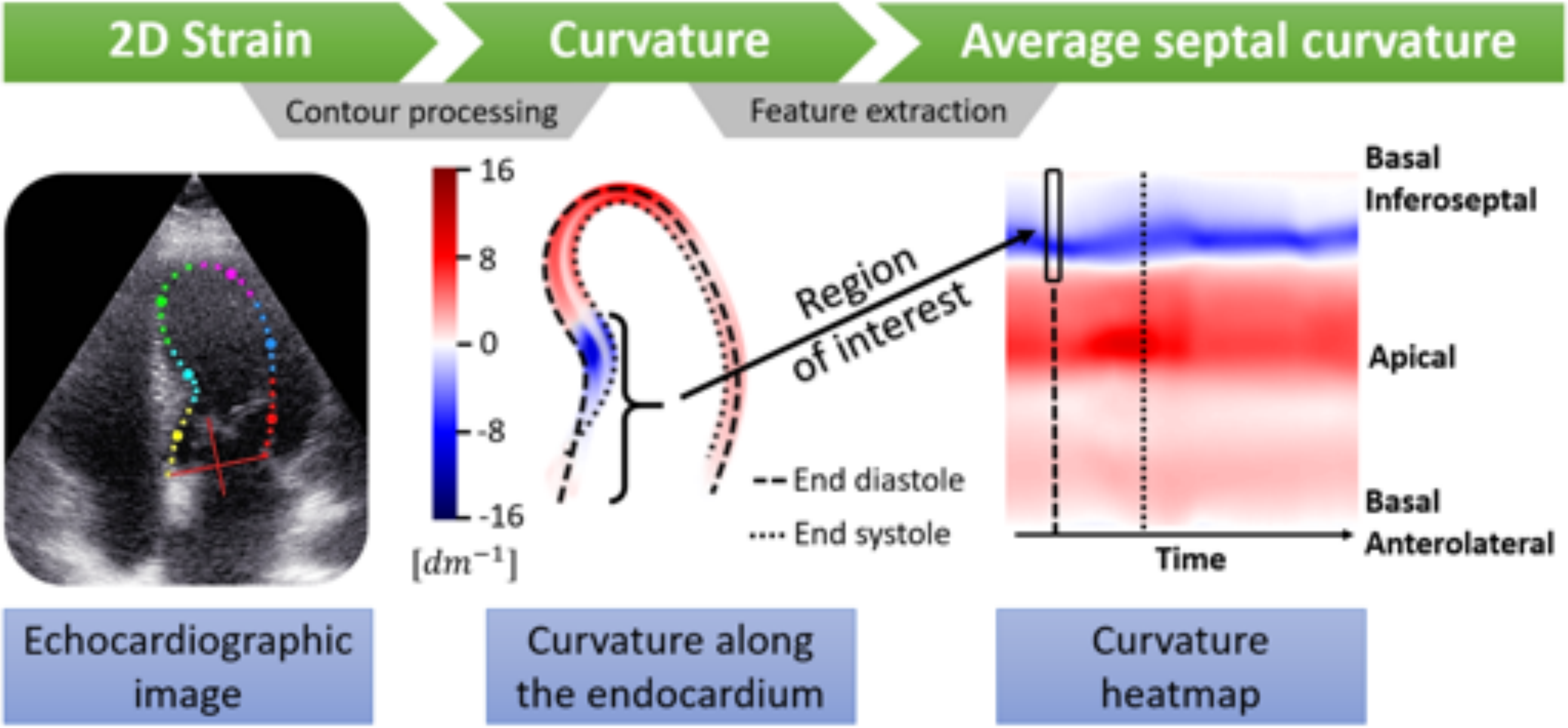


$$\mathcal{K}_i = \frac{|\dot{x}_i\ddot{y}_i - \dot{y}_i\ddot{x}_i|}{({\dot{x}_i}^2+{\dot{y}_i}^2)^{\frac{3}{2}}}$$

$$i \in \{1, 2, 3, \dots, N\}$$

The contour is extracted and processed to find a heatmap of curvature over time.

III: Curvature as a marker of hypertrophy



PIC GE Healthcare Universitetet i Oslo

Average septal curvature is much more strongly correlated with functional metrics associated with basal septal hypertrophy than traditional markers.

III: Curvature as a marker of hypertrophy

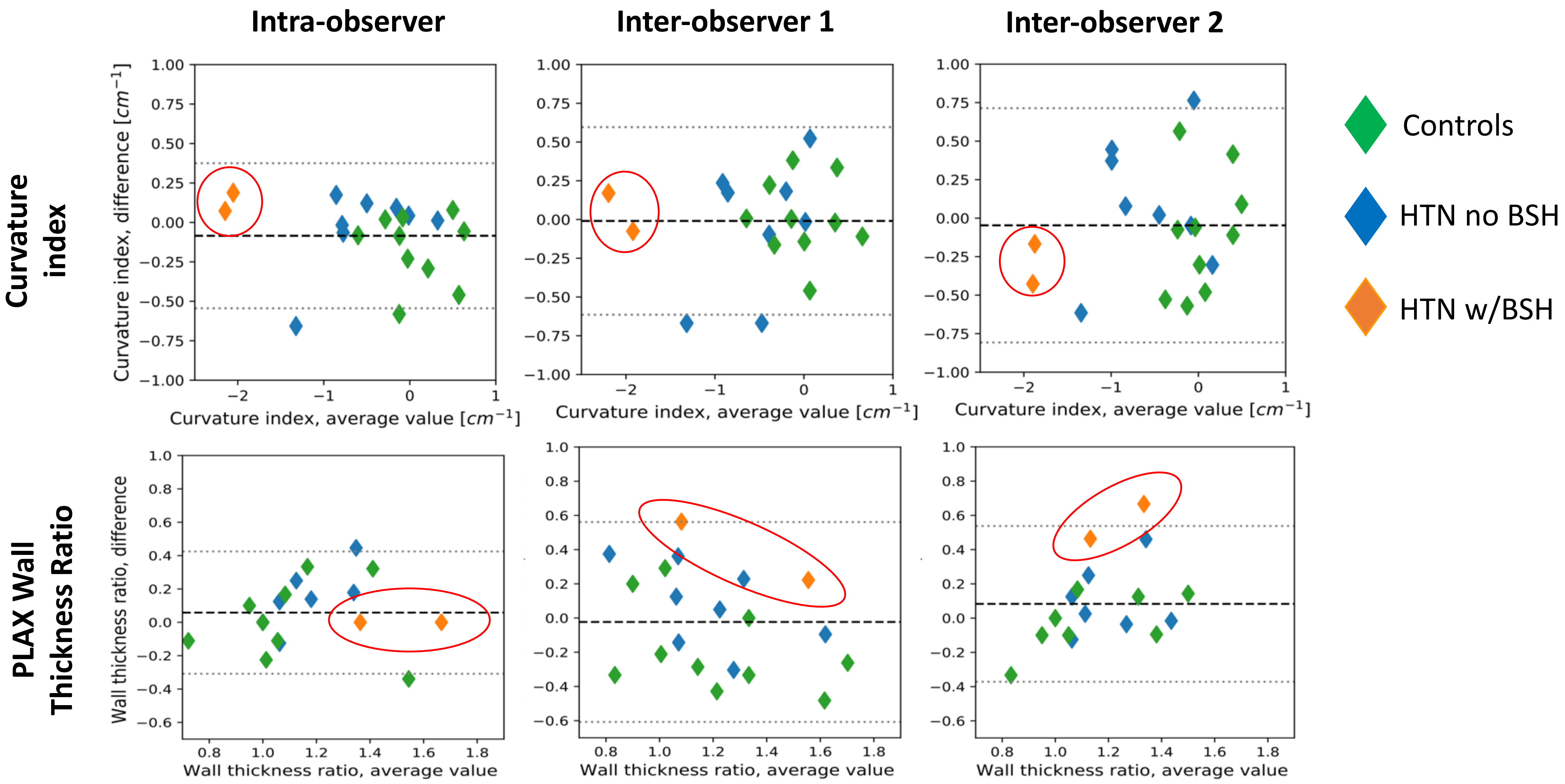
*significantly correlated

Variables	Spearman ρ – rank correlation		
	Average Septal Curvature	4CH Wall Thickness Ratio	PLAX Wall Thickness Ratio
Mitral annulus e' medial velocity	**0.234***	-0.156*	-0.158*
Mitral annulus a' medial velocity	-0.192*	**0.227***	0.163*
Mitral annulus e' lateral velocity	0.062	-0.116	-0.085
Average mitral annulus e' velocity	0.129	0.148	0.137
Basal septal strain	**-0.417***	0.341*	0.24*
Mid septal strain	**-0.164***	0.100	0.109
LA contractile strain	**-0.219***	0.158*	0.152
LA conduit strain	**0.159***	-0.137	-0.060
LA conduit/contractile strain ratio	**-0.232***	0.203*	0.140

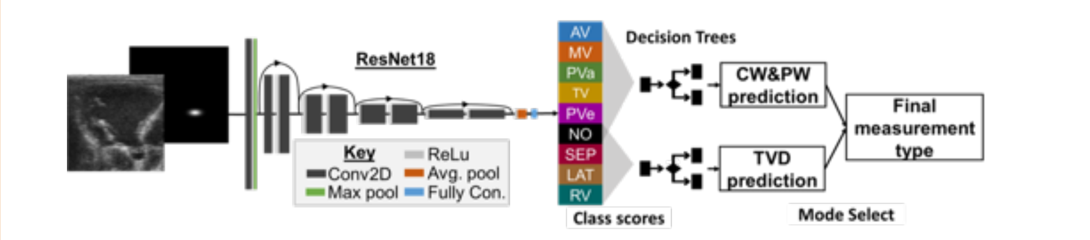
GE Healthcare Universitetet i Oslo

Average septal curvature is more reproducible.

III: Curvature as a marker of hypertrophy

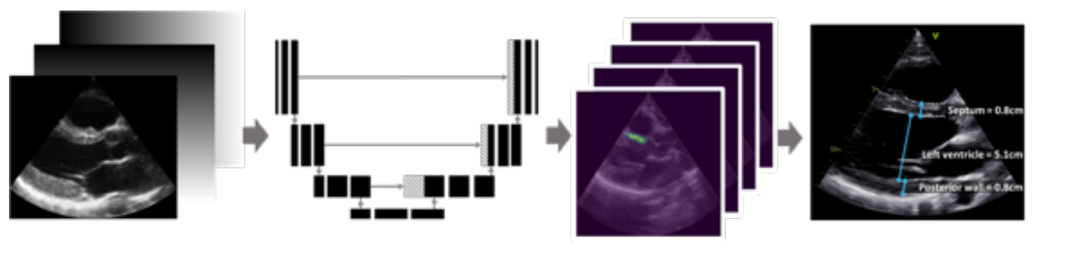


I: Doppler spectrum classification



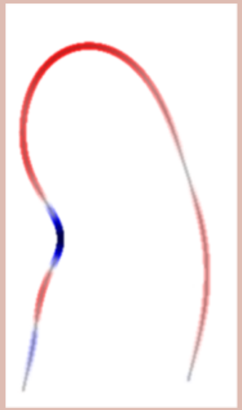
Highly accurate Doppler spectrum classification using **heatmap-encoding**, **multi-head training**, and **confidence metrics**.

II: Automated left ventricle dimension measurement



Accurate two-dimensional measurements of the left ventricle using **coordinate convolution**, **anatomical heatmaps**, and a multi-component loss function.

III: Curvature as a marker of basal septal hypertrophy



Curvature is a **marker** of basal septal hypertrophy. Compared to traditional metrics it:

- **Better correlates** to functional metrics
- Is more **reproducible**

Curvature can be **easily extracted** from existing tools and adds value to clinical workflows.

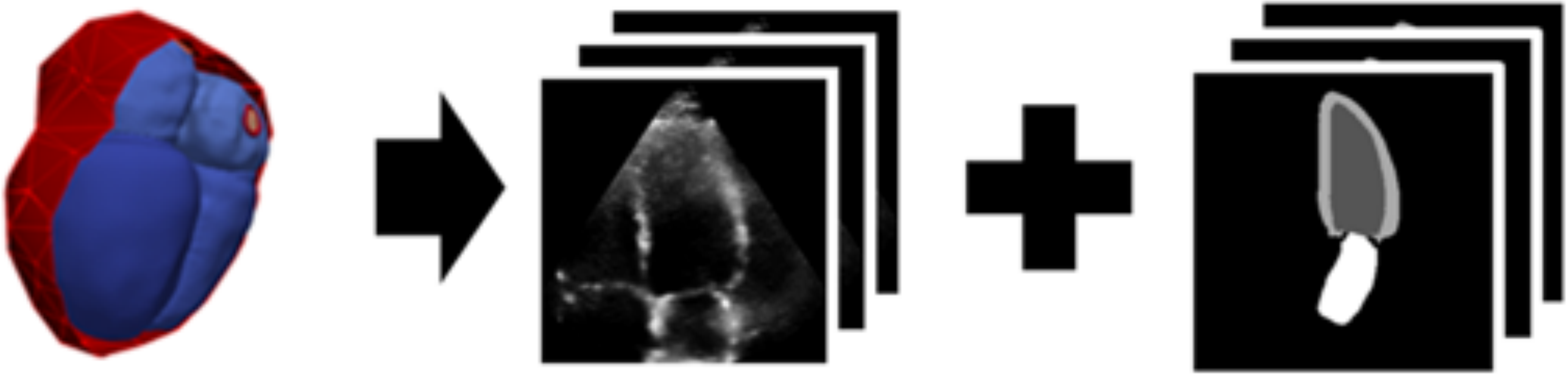
IV: Synthetic data generation

IV: Synthetic data generation

Gilbert, A., Marciniak, M., Rodero, C., Lamata, P., Samset, E., and McLeod, K. "Generating Synthetic Labeled Data from Existing Anatomical Models: An Example with Echocardiography Segmentation". In: *Transactions in Medical Imaging*. (2021).

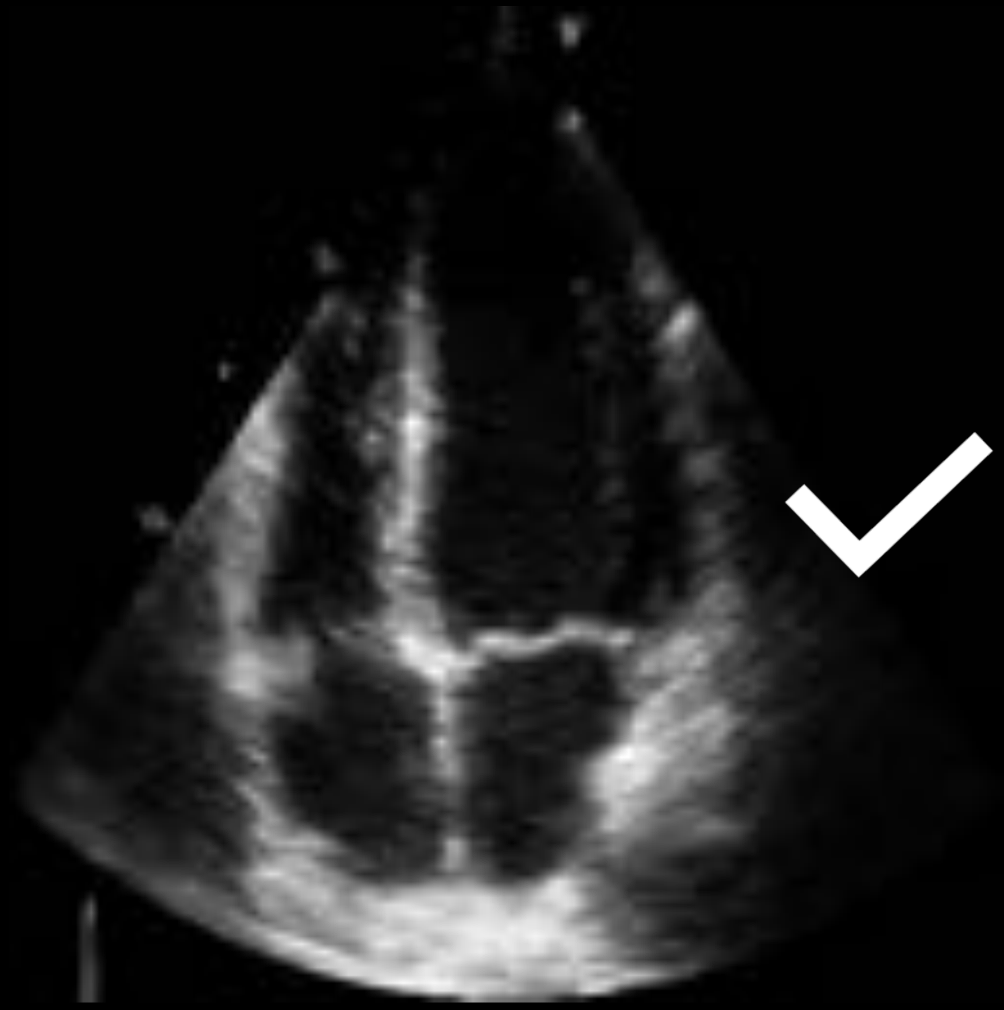
Goal: Speed up training of AI algorithms by generating synthetic images for training.

IV: Synthetic data generation



In general, deep learning algorithms are only as good as the data they are trained on.

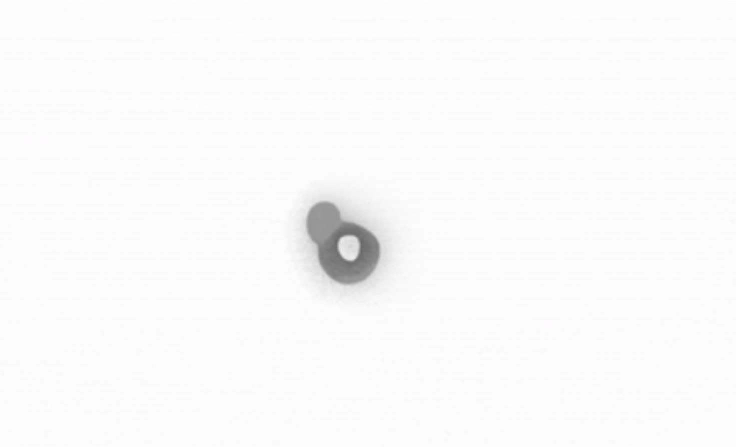
IV: Synthetic data generation



Generating high quality annotations in echo is difficult and time consuming.

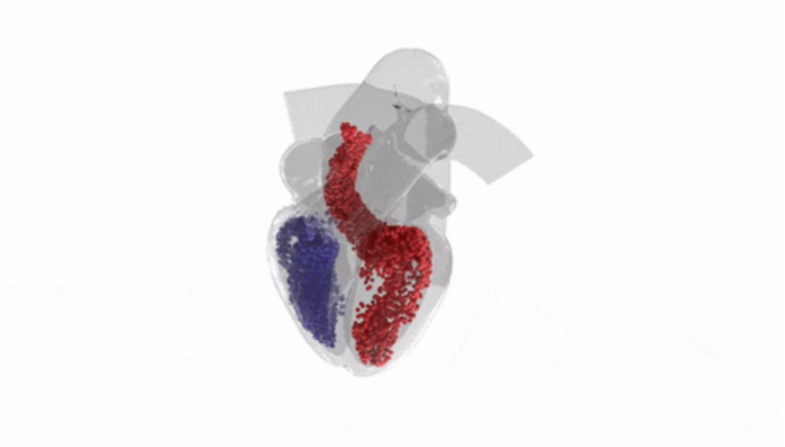
Computational models have been previously developed for research in a variety of areas.

IV: Synthetic data generation



Shape analysis

Electromechanical simulations



Fluid dynamics

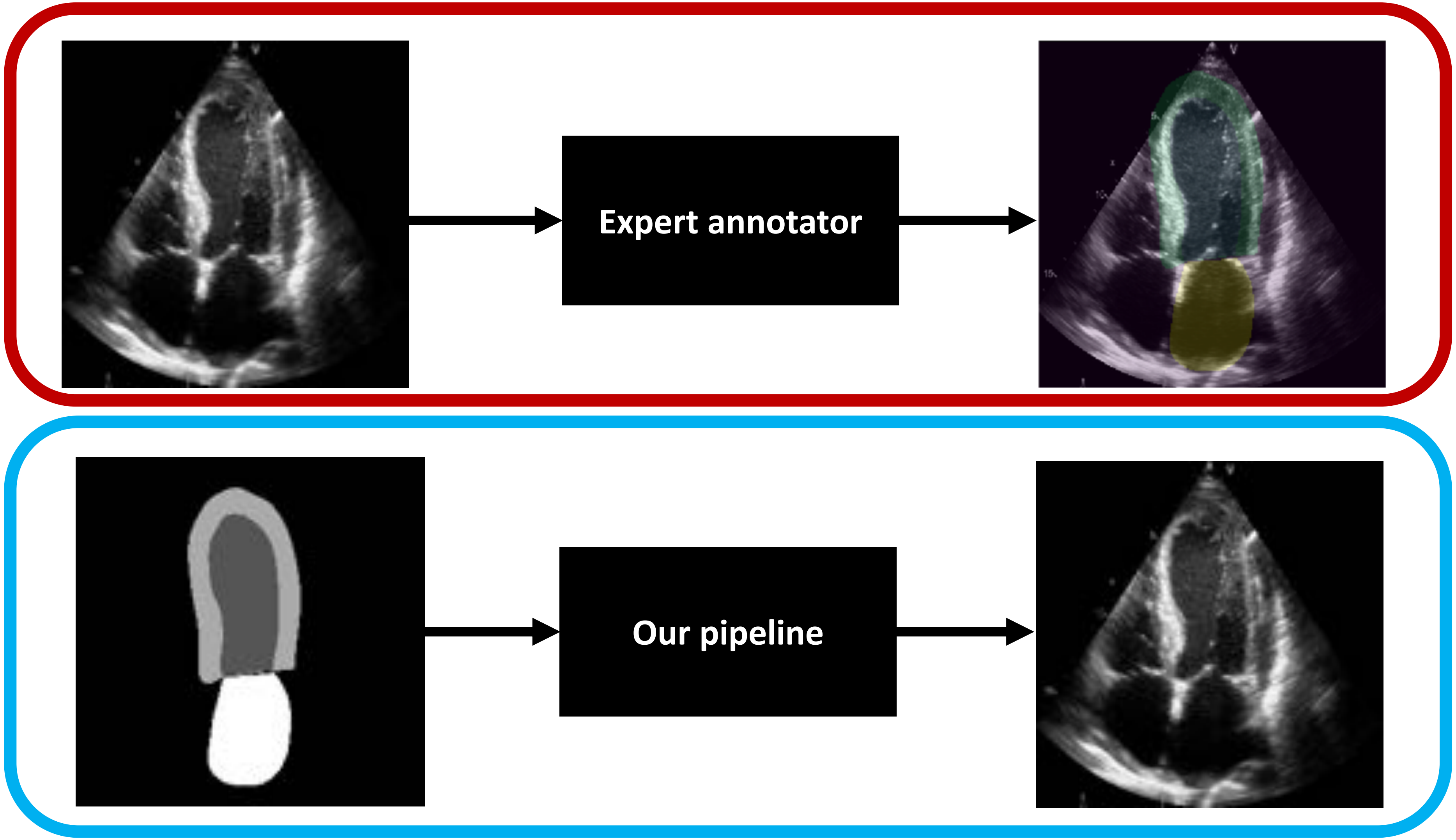
Let's use these models as sources of annotations

IV: Synthetic data generation

Instead of relying on annotators, our pipeline automatically generates images to match existing high-quality annotations.

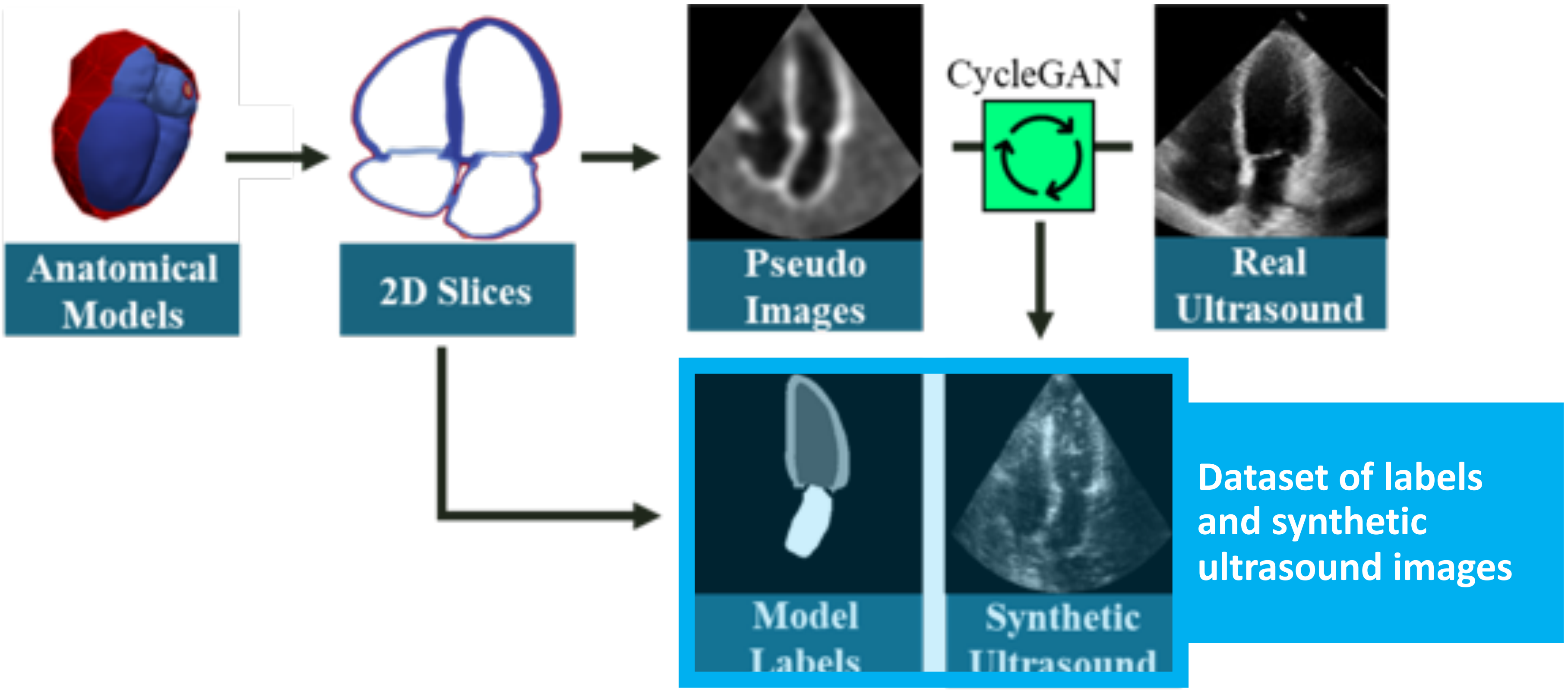
Manual workflow

Our workflow



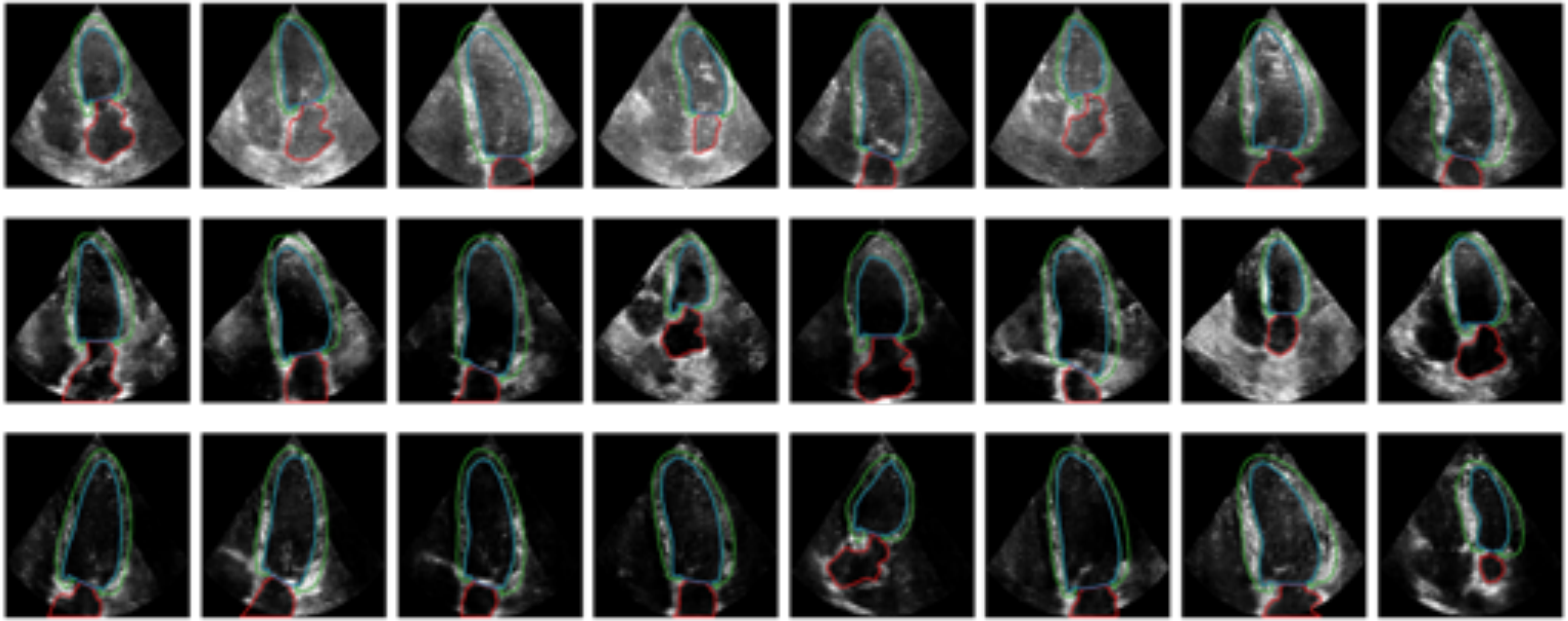
Our pipeline moves from anatomical models to a dataset of synthetic ultrasound images with paired labels.

IV: Synthetic data generation



Synthetic images appear realistic and contain accurate annotations that match structures.

IV: Synthetic data generation

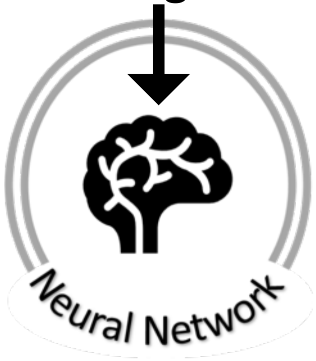


PIC GE Healthcare Universitetet i Oslo

Networks trained with synthetic data achieve accurate segmentations on real data

Real data

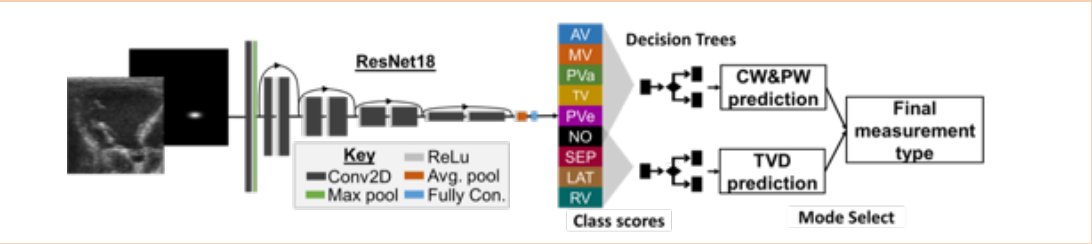
Real / synthetic
training data



Segmentation

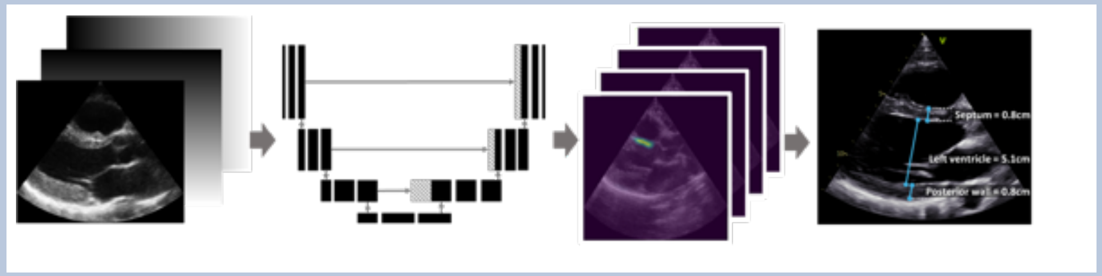
Method	LV_{endo} Dice score
Training/testing with **real data** *same dataset*	93.7
Training/testing with **real data** *different datasets*	89.1
Training with **synthetic data** Testing with **real data**	89.1
Inter-observer	87.9

I: Doppler spectrum classification



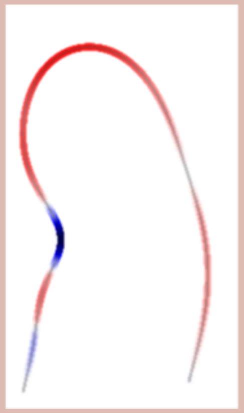
Highly accurate Doppler spectrum classification using **heatmap-encoding**, **multi-head training**, and **confidence metrics**.

II: Automated left ventricle dimension measurement



Accurate two-dimensional measurements of the left ventricle using **coordinate convolution**, **anatomical heatmaps**, and a multi-component loss function.

III: Curvature as a marker of basal septal hypertrophy

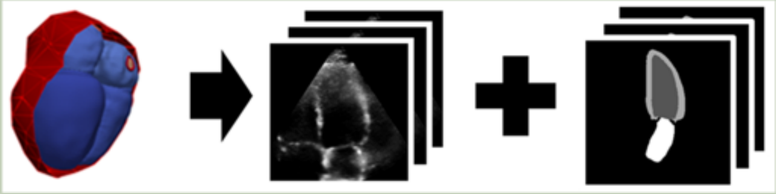


Curvature is a **marker** of basal septal hypertrophy. Compared to traditional metrics it:

- **Better correlates** to functional metrics
- Is more **reproducible**

Curvature can be **easily extracted** from existing tools and adds value to clinical workflows.

IV: Synthetic data generation



Proposed a method to **automatically generate** echo images with **a known ground truth** to facilitate training of deep learning algorithms.

High performance when using synthetic images

Summary

Goal: Make echocardiography more **efficient**, **reliable**, and **accurate**.

Efficiency:

- Automated classification and measurement workflows ***make analysis faster***.
- Automated generation of synthetic data ***increases speed*** of new tool development.

I II III IV

Reliability:

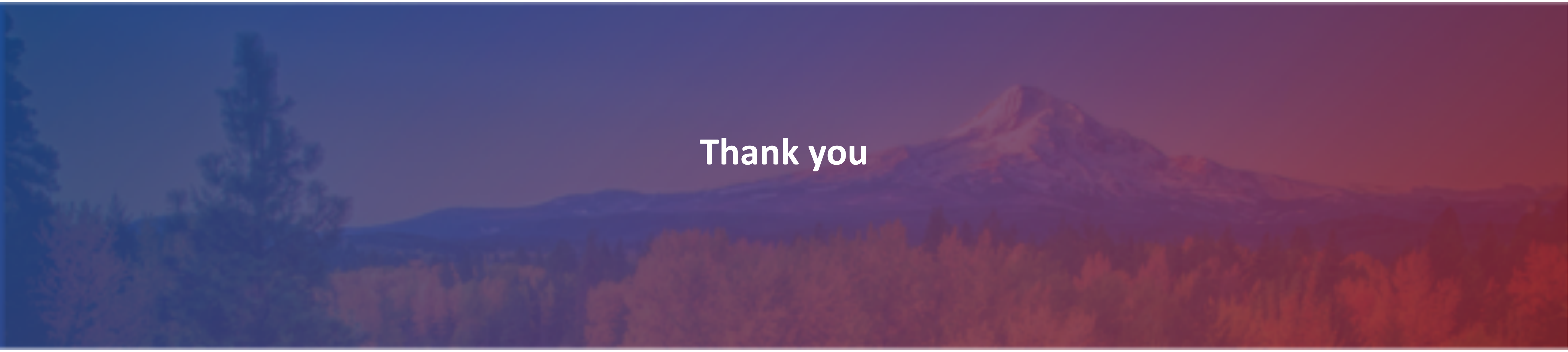
- Automated measurement workflows and new markers ***decrease variability***.
- Synthetic data ***increases consistency*** of labels.

I II III IV

Accuracy:

- New markers ***better correlate to*** functional metrics.
- Automated measurement workflow ***approaches inter-observer error***.

I II III IV



Thank you

This project has received funding from the European Union’s Horizon 2020 research and Innovation program under the Marie Sklodowska-Curie grant agreement No 764738