

Supplement:

Preprocessing

As a pre-processing step, echocardiogram videos that are clinically stored in the Dicom format were first processed automatically by extracting frames in the video. Once all the frames of various aspect ratios and image sizes were extracted, the frames were then converted to grayscale images and center cropped by removing the top and bottom 10% and the left and right 25% of the image. Once the images were cropped, they were matched to the average histogram of the training dataset and resized to the input of the model.

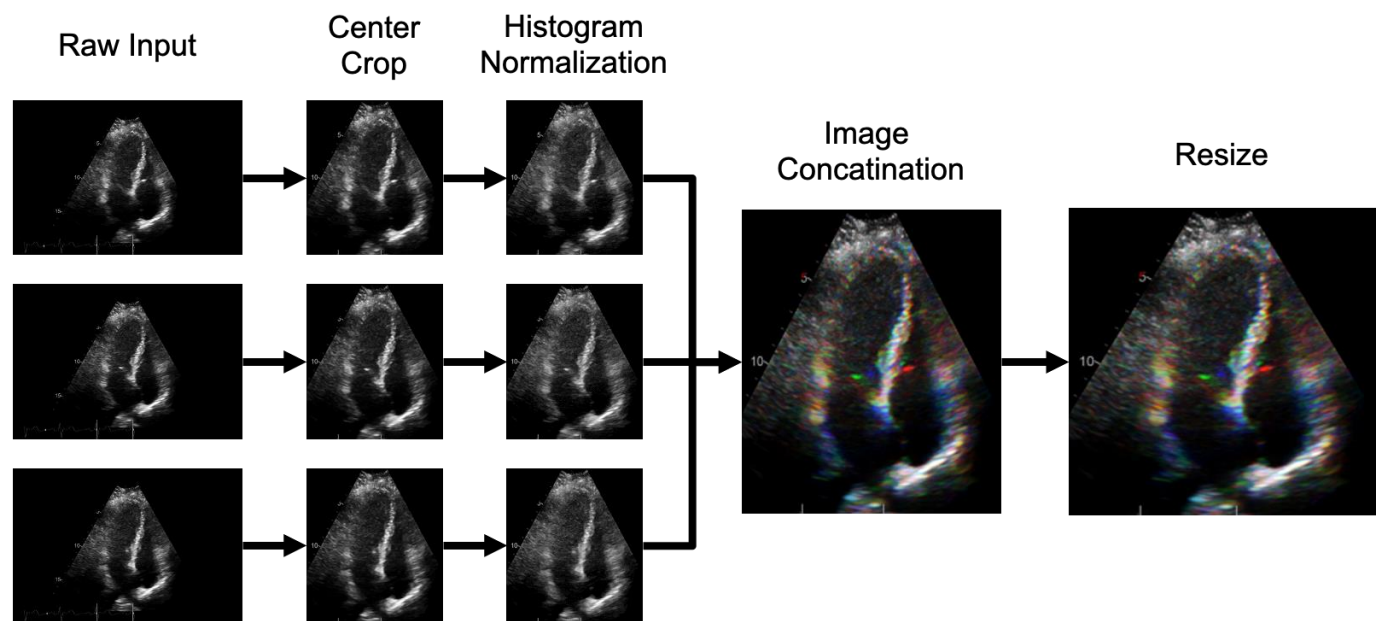
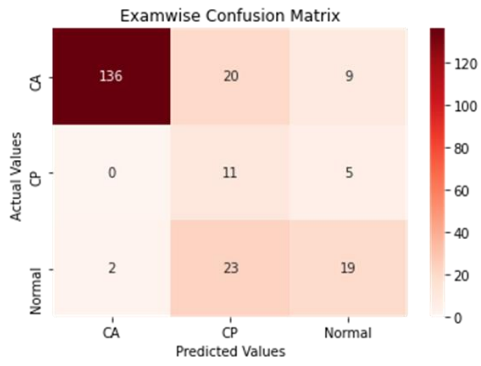
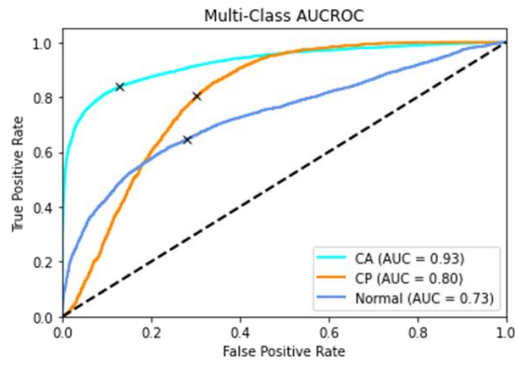
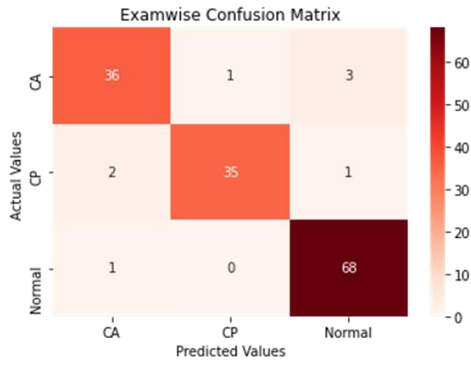
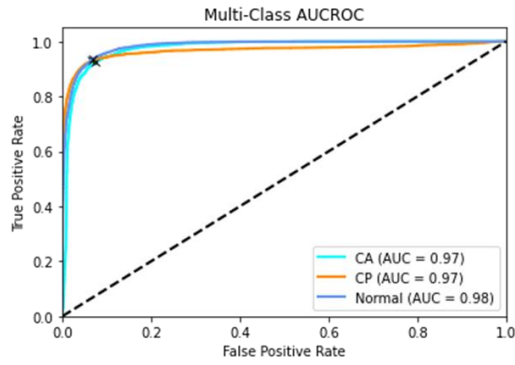


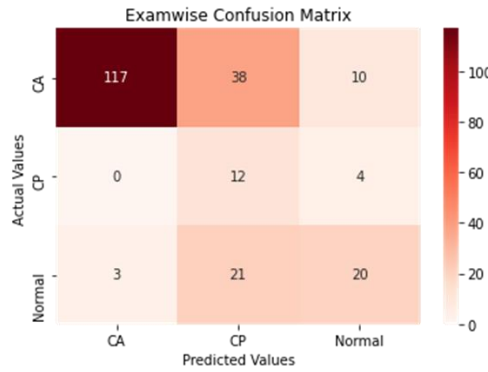
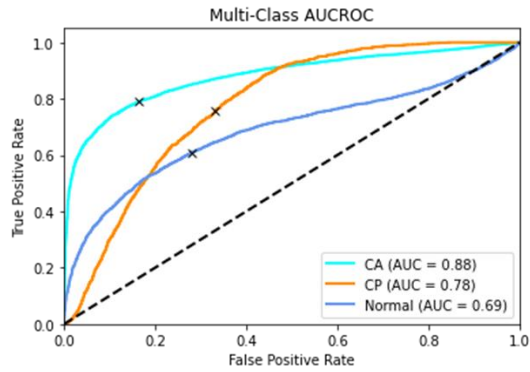
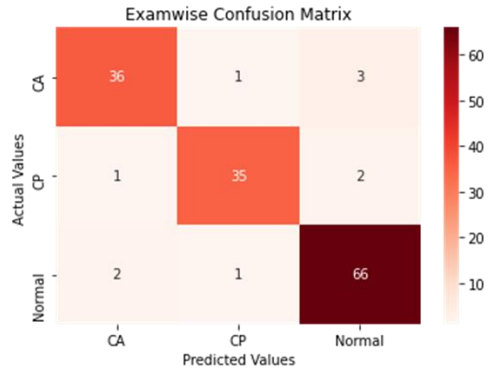
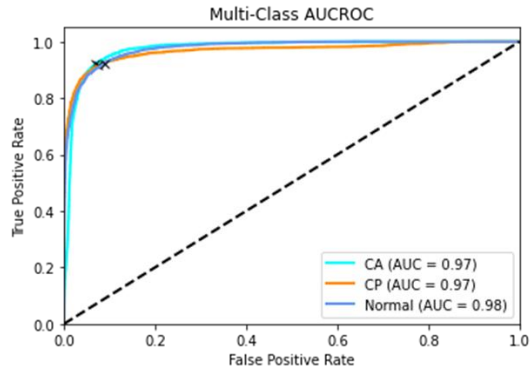
Figure Supp.1 . Preprocessing echocardiogram frames. Each image shows the output of the processing, and text above shows the process names. The red bars were used to anonymize the patient information for the figure.

Results: AUCROC and Confusion Matrix

ResNet50 figures



ResNeXt101 Results



EfficientNet-b2 Results

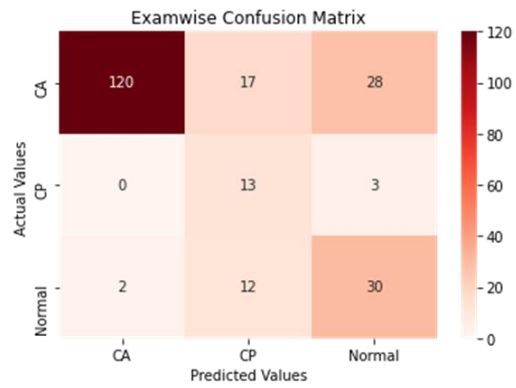
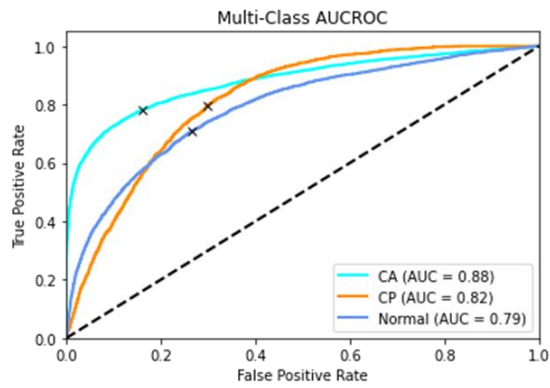
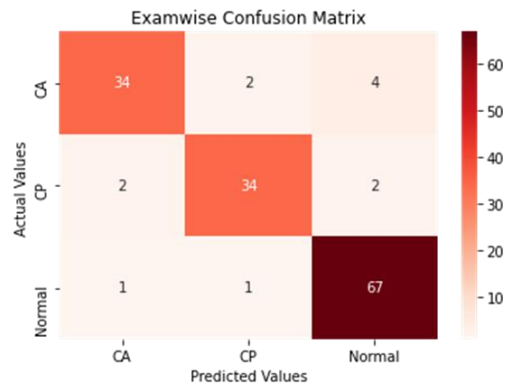
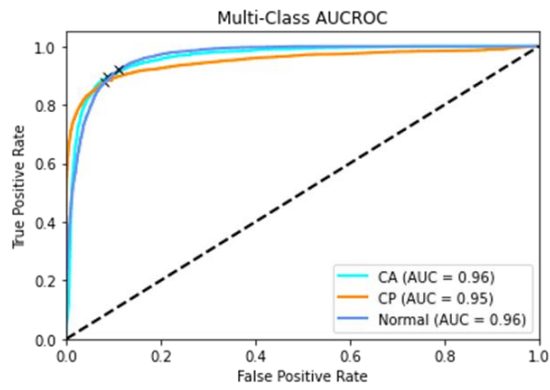


Figure Supp.2 . Comparative model performance on the internal (top) and external (bottom) dataset – AUROC and confusion matrix.

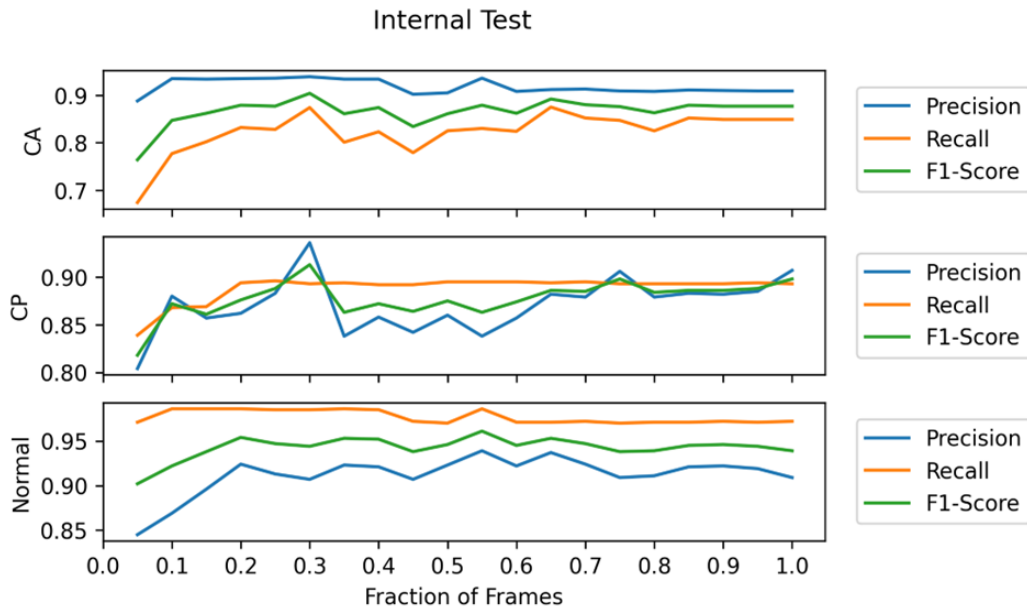
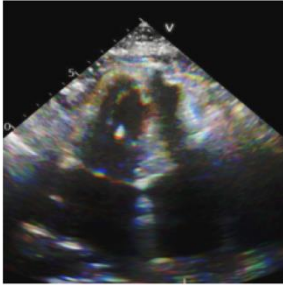
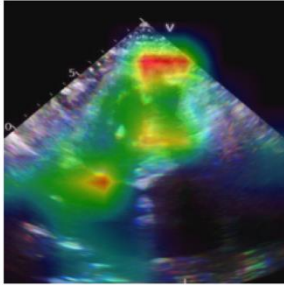
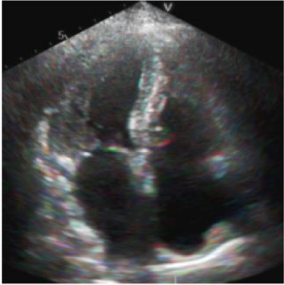
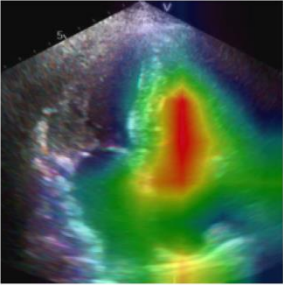
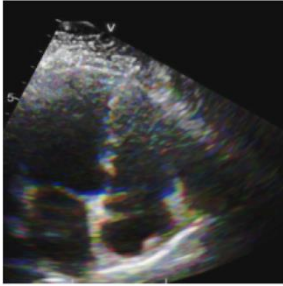
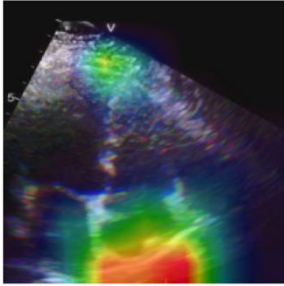
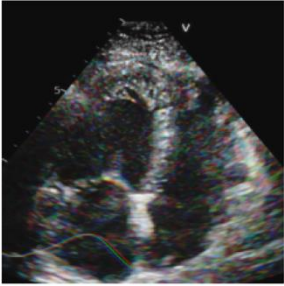
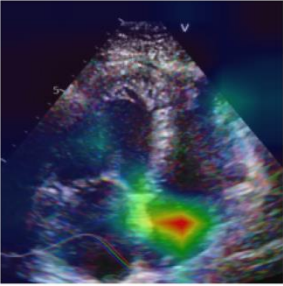


Figure Supp.3 . Frame-wise ablation – x-axis shows the fraction of frames and y axis shows the performance metrics.

Supplemental Table 1:

					
(a)		(b)			
CA: 0.044	CP: 0.859	Normal: 0.097	CA: 0.546	CP: 0.309	Normal: 0.145
					
(c)		(d)			
CA: 0.062	CP: 0.435	Normal: 0.503	CA: 0.271	CP: 0.168	Normal: 0.561

Supplemental Table 2: Gray-zone prediction probabilities of the figures in Figure 5.

Image	CA	CP	Normal
(a)	0.044	0.859	0.097
(b)	0.546	0.309	0.145
(c)	0.062	0.435	0.503
(d)	0.271	0.168	0.561