

1 **Development and validation of an end-to-end deep learning pipeline to measure**
2 **pericardial effusion in echocardiography**

3

4 Short title: Deep learning for pericardial effusion in echocardiography

5

6 Cheng-Ching Wu^{1,2,3}, Chi-Yung Cheng^{4,5}, Huang-Chung Chen⁶, Chun-Huei Hung⁷,

7 Tien-Yu Chen⁶, Chun-Hung Richard Lin⁴, I-Min Chiu^{4,5}

8 1. Division of Cardiology, Department of Internal Medicine, E-Da Hospital, Kaohsiung,

9 Taiwan

10 2. Division of Cardiology, Department of Internal Medicine, E-Da Cancer Hospital,

11 Kaohsiung, Taiwan

12 3. School of Medicine, College of Medicine, I-Shou University, Kaohsiung, Taiwan

13 4. Department of Computer Science and Engineering, National Sun Yat-sen

14 University, Kaohsiung, Taiwan

15 5. Department of Emergency Medicine, Kaohsiung Chang Gung Memorial Hospital

16 6. Division of Cardiology, Department of Internal Medicine, Kaohsiung Chang Gung

17 Memorial Hospital, Kaohsiung, Taiwan

18 7. Skysource Technologies Co., Ltd.

19

20 Correspondence: I-Min Chiu, ray1985@cgmh.org.tw

21 *Cheng-Ching Wu and Chi-Yung Cheng contributed equally to this work

22 Word count: 4426

23

24 **Abstract**

25 **Introduction**

26 Cardiac tamponade, caused by pericardial effusion (PE), is a life-threatening condition
27 that can be resolved by timely pericardiocentesis. Nevertheless, PE measurement
28 remains operator-dependent and may be difficult in some circumstances. Our study
29 aimed to develop a deep-learning pipeline that measures the amount of PE based on
30 raw echocardiography clips.

31 **Methods**

32 Echocardiographic examination data were collected from one medical center in
33 southern Taiwan from 2010–2018. Four commonly used cardiac windows, including
34 the parasternal long-axis, parasternal short-axis, apical four-chamber, and subcostal
35 views from included ultrasound examinations, were used for analysis. We proposed a
36 deep learning pipeline consisting of three steps: moving window view selection,
37 automated segmentation, and width calculation from a segmented mask. The pipeline
38 was then prospectively validated from 2019–2020 using a dataset from the same
39 hospital, and externally validated using data from another medical center in Taiwan.
40 Model performance was evaluated using mean absolute error, intraclass correlation
41 coefficient (ICC), and R-squared value between the ground truth and predictions.

42 **Results**

43 In this study, 995 echocardiographic examinations were included. Among these, 155
44 were used for internal validation and 258 were used for external validation. The
45 proposed pipeline had a predictive performance of ICC=0.867 for internal validation
46 and ICC=0.801 for external validation. It accurately detected PE with an area under the
47 receiving operating characteristic curve (AUC) of 0.926 (0.902–0.951) for internal
48 validation and 0.842 (0.794–0.889) for external validation. Regarding the recognition
49 of moderate PE or worse, the AUC values improved to 0.941 (0.923–0.960) and 0.907
50 (0.876–0.943) for internal and external validation, respectively. Of all the selected
51 cardiac windows, our model had the best prediction in the parasternal long-axis and
52 apical four-chamber views.

53 **Conclusions**

54 The machine-learning pipeline could automatically calculate the width of the PE from
55 raw ultrasound clips. The novel concepts of moving window view selection for image
56 quality control and computer vision techniques for maximal PE width calculation seem
57 useful in the field of ultrasound.

58

59

60 **Abbreviations**

61 A4C, apical four-chamber; AI, artificial intelligence; AUC, area under the receiver
62 operating characteristics curve; CGMH, Chang Gung Memorial Hospital; CNN,
63 convolutional neural network ; DICOM, Digital Imaging and Communications in
64 Medicine; EDH, E-Da Hospital; GPU, graphics processing unit; ICC, intraclass
65 correlation coefficient; MWVS, moving window view selection; PE, pleural effusion;
66 PLAX, parasternal long-axis; PSAX, parasternal short-axis; SC, subcosta

67

68 **Introduction**

69 Pericardial effusion (PE) is an acute or chronic accumulation of fluid within the
70 pericardial space which is usually revealed by transthoracic echocardiography. Fluid
71 accumulation increases pressure in the pericardial sac, leading to compression of the
72 heart and subsequent cardiac tamponade. In acute settings, only 100–150 mL of fluid is
73 necessary for cardiac tamponade to occur, which results in impaired diastolic filling
74 and reduced cardiac output. PE is a life-threatening issue that can be resolved by timely
75 pericardiocentesis.¹ Therefore, the early detection of PE and measurement of PE width
76 are important.

77 Echocardiography remains the gold standard imaging modality for verifying the
78 presence of PE by demonstrating fluid collection in the pericardial space.^{2,3} It is
79 considered the first-line imaging choice since it is less costly, more portable, more
80 widely available, and can provide a comprehensive anatomical and functional
81 assessment compared with computed tomography and magnetic resonance imaging.^{4,5}
82 Nevertheless, the presence and grading of PE is associated with intraoperative
83 uncertainty. For example, mild PE does not always correspond to true effusion and can
84 be indicative of pericardial fat.⁶ Furthermore, the image quality of a transthoracic
85 echocardiogram might be compromised by elements such as the female breast or
86 obscuration by bone or lung.⁴ Early disclosure of the precise PE grade might be difficult
87 in some circumstances and is dependent on operator experience.

88 Artificial intelligence (AI) has been used in many clinical settings to assist in the
89 diagnosis of conditions based on echocardiograms. Considerable effort has been
90 devoted to themes such as left ventricular function assessment, regional wall motion
91 abnormality, right ventricular function, valvular heart disease, cardiomyopathy, and
92 intracardiac mass.⁷⁻¹⁰ Regarding the diagnosis of PE, one study in 2020 that used a deep
93 learning model to detect PE in echocardiography achieved an accuracy of 0.87–0.9.¹¹ In
94 daily practice, further information on PE width and severity is critical for initializing
95 the necessary interventions. To the best of our knowledge, no study has analyzed the

96 grading of PE via machine learning. Thus, our study attempted to develop a deep
97 learning model using echocardiography for PE detection and PE width measurement.
98 In addition, to better deploy the deep learning model, we propose an end-to-end
99 guideline that can output the prediction results from raw ultrasound files.

100

101 **Method**

102 The data collection and protocols utilized in this study were authorized by the
103 Institutional Review Board of E-Da Hospital (EDH; no: EMRP24110N) and the
104 Institutional Review Board of Kaohsiung Chang Gung Memorial Hospital (CGMH; no:
105 20211889B0 and 202101662B0).

106 *Data Collection*

107 In this study, images from routine echocardiography were generated at two medical
108 centers, EDH and CGMH, in southern Taiwan. The deep learning model was trained
109 and internally validated in EDH and externally validated in CGMH.

110 During data collection, we used “pericardial effusion” as a keyword to search the
111 Hybrid picturE Report System in EDH to collect the examination list. We obtained
112 patients’ raw data from transthoracic echocardiography examinations with PE

113 performed at EDH between January 1, 2010, and June 30, 2020. These data were
114 divided into training and validation datasets based on the respective index dates of the
115 examinations. Examinations with index dates prior to December 31, 2018, were used
116 for the development of the model, and examinations with index dates after January 1,
117 2019, were used for internal validation. To test the generalizability of the model, we
118 retrieved echocardiography data from CGMH between January 1, 2019, and June 30,
119 2020, for external validation. The study flowchart and data summary are presented in
120 Figure 1.

121

122 *Echocardiography*

123 Images were gathered in a normal manner, with patients lying in the left lateral
124 decubitus position. The ultrasound system (IE33, Philips Healthcare; S70, GE
125 Healthcare; or SC2000, Siemens Healthineers) was used to perform echocardiographic
126 examinations in EDH. Data from CGMH for external validation were acquired using
127 EPIC7 (Philips Healthcare), Vivid E9 (GE Healthcare), or SC2000 (Siemens
128 Healthineers). All examinations were saved in picture archiving and communication
129 systems in the Digital Imaging and Communications in Medicine (DICOM) format.

130 After extracting the raw DICOM files, we processed the image from each patient to
131 select the proper echocardiography views for developing a deep learning pipeline. The
132 selected views were the parasternal long-axis (PLAX), parasternal short-axis (PSAX),
133 apical four-chamber (A4C), and subcostal (SC) views. The ground truth of PE width
134 was annotated from the examination report by a proficient cardiac physiologist and was
135 inspected by a cardiologist who delegated the confirmed reports.

136

137 *Deep Learning Model Development*

138 In this study, we developed an end-to-end pipeline for the automated measurement of
139 PE based on the steps outlined below (Figure 2). The training subset of videos from
140 EDH was used for the three main tasks of our pipeline:

141

142 *Step one: moving window view selection (MWVS)*

143 We proposed a pipeline to directly manage echocardiography files from the
144 workstation, similar to the work done by Zhang et al. and Huang et al., with some
145 adjustments (Figure 2A).^{10,12} To distinguish the four primary views (PLAX, PSAX,
146 A4C, and SC) from other views during each examination, we developed the first deep
147 neural network model. This model was a ResNet-50-based two-dimensional model that
148 aimed to classify each frame from the extracted DICOM files of echocardiography into

149 the selected view types.¹³ To train this model, we randomly selected 6,434 images from
150 the training dataset of EDH and labeled them according to the four primary views or
151 other views, including low-quality views. We trained the model with data splitting of
152 80% and 20% for the training and validation sets, respectively. The model weight with
153 the best prediction performance in the validation set during the training process was
154 preserved. The prediction accuracy was assessed for each view class and weighted
155 average result.

156 While managing the input video from the patient, we used 48 frames moving window to
157 filter all videos. For each video, we retrieved a clip of 48 frames with the best
158 confidence with regard to the specific view type by majority voting (Figure 3). The
159 MWVS concept was used not only as a view classifier but also for quality control. This
160 process not only helps the algorithm to identify the right video but also retrieves the
161 best 48 frames of the video with regard to image quality. If a video did not contain any
162 of the 48 consecutive frames that consisted of qualified frames higher than 50% from
163 one of the four primary views, it was excluded from further analysis. Moreover, the
164 view-classifying confidence levels for all images obtained from the clip of 48 frames
165 were averaged to evaluate the general image quality, and they correlated with further
166 performance. Ultrasound videos with an average confidence level of < 0.8 were
167 excluded from further automated segmentation.

168

169 *Step two: automated segmentation*

170 From the dataset, we annotated 2,548 randomly selected frames in the EDH training
171 dataset which were evenly distributed over the four primary views. For each view, we
172 manually labeled the segmented area for PE at three different phases in the cardiac
173 cycle: end-systolic phase, end-diastolic phase, and middle phase between the two
174 aforementioned phases. We also labeled the segmented areas for the four cardiac
175 chambers to enhance the model performance in separating these fluid-containing areas.
176 We used a mask region-convolutional neural network (R-CNN) as the framework to
177 train object instance segmentation based on the labeled ground truth (Figure 2B). The
178 model was trained with 80% and 20% data splitting for the training and validation sets,
179 respectively. Mask R-CNN is commonly used for instance segmentation tasks in
180 medical applications because it can simultaneously perform pixel-level segmentation
181 and classification of multiple target lesions.¹⁴ The implemented model generates
182 bounding boxes and targeting masks for each instance of an object in an image. As such,
183 the input comprised consecutive ultrasound frames, and the output comprised a
184 segmented mask, which indicated the corresponding four cardiac chambers and PE.
185 The accuracy of the segmentation model was assessed using the Dice coefficient
186 metric.

187

188 *Step three: measurement of pericardial effusion*

189 After generating a segmented mask for PE, we proposed a computer vision technique
190 (maximal width calculator of the segmented mask) to calculate the largest width of the
191 PE in each ultrasound frame (Figure 2C). For this task, we iterated through the vertical
192 axis in each frame and hypothetically drew a horizontal line to see if there was any
193 intersection between the segmented mask and horizontal line. If an intersection existed,
194 we obtained a normal line from the edge of the mask over the intersection point. The
195 length of the normal line that passes through the segmented mask was counted as the
196 width of the PE at that intersection point. The largest width of the PE through the
197 iteration over the vertical axis was regarded as the width of the PE of the frame. The
198 same technique was applied to all 48 frames in the ultrasound video to provide the
199 optimal PE width.

200

201 *Statistical Analysis*

202 Continuous variables are presented as mean (standard deviation) if normally distributed,
203 otherwise they are presented as median with the interquartile range. Dichotomous data
204 are presented as numbers (percentage). Categorical variables were analyzed using the

205 χ^2 test. Continuous variables were analyzed using the independent-sample t-test if
206 normally distributed; otherwise, the Mann-Whitney U test was used.

207 Model performance regarding PE width measurement was analyzed based on the mean
208 absolute error, intraclass correlation coefficient (ICC), and R-square value between the
209 ground truth and prediction. We further examined the prediction of the existence of PE
210 and moderate PE using sensitivity, specificity, and area under the receiver operating
211 characteristic curve (AUC). The deep learning models in the proposed pipeline were
212 developed using the TensorFlow Python package. Image manipulation was performed
213 using OpenCV 3.0 and scikit images. All analyses were performed using SPSS for
214 MAC version 26.

215

216 **Results**

217 In this study, 737 examinations from EDH were included in the analysis, of which 582
218 were included in the training set and 155 were in the internal validation set. For the
219 external validation, 258 examinations from CGMH were included. Because there were
220 less than 10 SC views in the CGMH dataset, the SC view was excluded from further
221 analysis in the external validation.

222 The demographic and clinical characteristics of those who underwent
223 echocardiography are presented in Table 1. The mean ages of patients in the training,

224 internal validation, and external validation set were 67.4 ± 15.4 , 59.8 ± 19.2 , and
225 66.4 ± 16.1 years, respectively; 46.5% and 63.2% of patients in the internal and external
226 validation groups, respectively, had PE. The average ejection fraction was $64.3\pm 7.1\%$
227 in the internal validation group and $61\pm 13.9\%$ in external validation group.

228 The view classifier achieved an average accuracy of 0.91 and 0.87 in predicting image
229 classes in the training and validation sets, respectively. The independent accuracies in
230 the validation set for each class were 0.90, 0.87, 0.93, 0.76, and 0.88 for PLAX, PSAX,
231 A4C, SC, and others, respectively.

232 After MWVS, most of the ultrasound videos, ranging from 80–100% among the four
233 selected ultrasound views in EDH, successfully passed through for the segmentation
234 model. In the external validation (CGMH dataset), 686 ultrasound videos from the four
235 selected views were obtained. Our MWVS scanned through all DICOM files of 258
236 patients, and 53.2% of the 686 ultrasound videos were preserved for segmentation
237 inference. The videos selected by our pipeline were further checked by a cardiologist,
238 and none of them were misclassified into other cardiac views.

239 In image segmentation, the mask R-CNN-based model effectively localized the cardiac
240 chambers and PE area within the four different views (Figure 4). In the validation set
241 consisting of 510 images, the average Dice coefficient ranged from 0.67–0.82 among
242 the four different views, with the SC view being the lowest. PE segmentation in the

243 PLAX view showed the best Dice result 0.72, while the SC view had the poorest result
244 (0.56) (Table 2). Based on the segmentation of the PE area, we calculated the maximal
245 PE width from each frame.

246 Figure 4 shows the scatter plot of the PE width measurement between the ground truth
247 and model prediction after finding the largest normal line passing the segmented mask
248 in each frame. Compared with the ground truth, we reported the absolute difference and
249 correlation between automated and manual measurements of PE in both the internal
250 (EDH) and external (CGMH) validation datasets. The mean absolute error was 0.33 cm
251 and 0.35 cm in internal and external datasets, respectively. Interobserver variability
252 was highly correlated for the measurement of PE width between our model and human
253 expert (ICC=0.867, $p<0.001$, EDH; ICC=0.801, $p<0.001$, CGMH). The R2 was 0.594
254 for EDH and 0.488 for CGMH validation dataset.

255 Our model accurately detected the existence of PE in the internal validation
256 (AUC=0.926 [0.902–0.951]) and external validation (AUC=0.842 [0.794–0.889]).
257 With regard to recognizing moderate PE or worse, the AUC values improved to 0.941
258 (0.923–0.960) and 0.907 (0.876–0.943) in the internal and external validation groups,
259 respectively.

260 We further performed a stratified analysis of the model prediction in the different
261 echocardiography views. In the internal validation, the model prediction of PE width

262 was highly correlated with the ground truth in the four different views, with ICC
263 ranging from 0.802–0.910. The PLAX and A4C views appeared to have the best
264 prediction results with ICCs of 0.910 (0.876–0.935) and 0.907 (0.871–0.932),
265 respectively. In the external validation, similar to internal validation, the model
266 performed better in the PLAX and A4C views, with ICCs of 0.807 (0.726–0.864) and
267 0.897 (0.846–0.931), respectively. The other performances are listed in Table 3.

268

269 **Discussion**

270 Computer vision and deep learning models have proven useful for aiding
271 echocardiography interpretation, estimating cardiac function, and identifying local
272 cardiac structures. In recent years, deep learning algorithms have also been applied to
273 facilitate the diagnosis of PE.¹⁵ Nayak et al. developed a CNN that detected PE in the
274 A4C and SC views with accuracies of 91% and 87%, respectively.¹¹ Pericardiocentesis
275 is an essential therapeutic procedure for the treatment of symptomatic PE. However,
276 this therapeutic procedure can be life-threatening. The major complications of
277 pericardiocentesis are mortality, cardiac arrest, cardiac perforation, and cardiac
278 chamber laceration, while other complications include pneumothorax, supraventricular
279 tachycardia, and pneumopericardial fistula.¹⁶ Blind pericardiocentesis is associated
280 with a mortality rate of 6% and complication rates of 20–50%^{17,18}. With

281 ultrasound-assisted pericardiocentesis, the mortality rate is <1%, and the overall
282 complication rate is approximately 4–20%.¹⁸⁻²⁰ Therefore, it is important to identify the
283 location and distribution of pericardial fluid while avoiding the accidental puncture of
284 vital organs.

285 In this study, we proposed a machine learning pipeline that could process raw DICOM
286 files from ultrasound and predict the PE width in clinical practice. This pipeline
287 combines two steps of the deep learning model and one technical calculation algorithm
288 to accurately predict PE width. Few efforts have been made to predict PE existence,¹¹
289 with some studies being based on computed tomography scans.^{21,22} To the best of our
290 knowledge, this is the first video-based machine learning model to predict the PE width
291 using echocardiography. The correlation between the prediction of our model and
292 human experts was highly in both the internal and external validation datasets, with the
293 best performance noted in the PLAX view. The speed of inference from accessing the
294 file to the output for our model in one graphics processing unit (GPU; NVIDIA RTX
295 3090) was approximately 30–40 s for one examination, which is usually faster than
296 human assessment.

297 The two methods proposed in this study are novel concepts, including MWVS for view
298 selection and the maximal width calculator of the segmented mask. These two methods

299 are crucial for real-world predictions, particularly for relatively smaller datasets. Many
300 previous studies used datasets manually selected by human experts during dataset
301 cleaning for machine learning and used only “textbook-quality” images for
302 training.²³⁻²⁶ In contrast, we hypothesized that an analytic pipeline could automatically
303 analyze echocardiograms and be easily applied to personal devices or web applications.
304 Hence, it is important to exclude processes that require an expert sonographer or a
305 cardiologist. Madani et al. trained a CNN to simultaneously classify 15 standard
306 echocardiogram views acquired based on a range of real-world clinical variations, and
307 the model showed high accuracy for view classification.¹⁹ Similarly, in our study, we
308 used echocardiogram video clips randomly obtained from the real world, which were
309 taken for a variety of clinical purposes, including ejection fraction calculation and for
310 detecting PE, valve disease, regional wall abnormality, cardiomyopathy, and
311 pulmonary arterial hypertension. We developed an initial screening model for view
312 classification and quality control. All raw images from the medical image database
313 were input into the screening model, leaving a specific view of sufficient quality for
314 diagnosis. In addition, with the “moving window” concept, we retrieved only clips with
315 48 consecutive frames that fulfilled the image quality. We found that ICC and
316 diagnostic accuracy were significantly improved after MWVS. By avoiding limited or

317 idealized training datasets, we believe that this model is broadly applicable to clinical
318 practice.

319 Rather than previous studies using three-dimensional CNN architecture for training the
320 view classifier,¹⁰ we used a two-dimensional ResNet structure. The two-dimensional
321 structure consumes significantly less computing resources and can be deployed as a
322 real-time feedback system for ultrasound operators using only one GPU. We proposed
323 MWVS by combining a two-dimensional image classifier with a moving-window
324 algorithm for clinical usage. MWVS is a novel concept that has not yet been proposed
325 in the field of echocardiography assisted by machine learning. MWVS plays the role of
326 an image quality filter, and the major function of MWVS is to ensure image quality in
327 keeping with the next step in the pipeline. In EDH, echocardiography is performed by
328 well-trained technicians who follow the protocol designed by the echocardiologist
329 consensus committee. Therefore, the original images from EDH had good homogeneity,
330 and MWVS filtered out fewer patients. At CGMH, echocardiography is performed by a
331 separate echocardiologist who has the respective habit to perform echo study.
332 Therefore, the original images from CGMH had poor homogeneity, and MWVS
333 filtered out more patients. This finding proved that MWVS plays a significant role in
334 maintaining image quality. However, this finding also confirms that the applicability of
335 machine learning depends on image homogeneity.

336 After segmenting the PE, we developed a computer vision technique to calculate the
337 largest PE width. The current categorization of PE size depends on linear
338 measurements of the largest width of the effusion at end-diastole and is graded as small
339 (<1 cm), moderate (1–2 cm), and large (>2 cm).²⁷ This semiquantitative classification
340 method is prone to errors because of the asymmetric loculated effusion and shifts in
341 fluid location during the cardiac cycle.²⁸ Therefore, an automated calculation system
342 could help identify the largest width of the PE in every ultrasound frame without any
343 errors. Compared with AI-based models, the computer vision technique is more similar
344 to the method used by human experts. The AI-based model not only consumes more
345 computing resources but also requires a large number of datasets for training and
346 validation. To our knowledge, our study is the first to use computer vision techniques
347 for classifying PE grades.

348 This study had certain limitations. First, we conducted the study retrospectively and
349 trained our model using only one hospital dataset. The sample size and ethnic diversity
350 were limited and the results may not be generalizable to other areas. The dataset should
351 have greater heterogeneity in a multicenter setting. However, the enrolled images were
352 obtained using different ultrasonography machines operated by several
353 echocardiographers. Moreover, we used standard views for diagnosis, and model
354 achievement was similar during external validation. Second, although we graded the

355 amount of PE, there was no information on whether there was a sign of cardiac
356 tamponade, because PE volume does not necessarily correlate with clinical
357 symptoms.²⁹ Further research should also evaluate the collapsibility of the cardiac
358 chambers and the presence of tamponade signs.

359

360 **Conclusion**

361 We developed a machine-learning pipeline that automatically calculates the width of
362 the PE from raw ultrasound clips. The model achieved high accuracy in detecting PE
363 and predicting the PE width in both internal and external validation. The concept of
364 MWVS for image quality and computer vision techniques for maximal PE width
365 calculators is a novel application in the field of ultrasound.

366 **Figure Legend**

367 **Figure 1. Flow chart illustrating the study design and data summary**

368 **Figure 2. Deep Learning Pipeline for PE Measurement**

369 This graphic demonstrated the end-to-end pipeline proposed in this study. The
370 pipeline consisted of three major steps, which were moving window view selection
371 (Figure 2A), automated PE segmentation (Figure 2B), and PE width calculation from

372 segmented mask (Figure 3B). The pipeline was able to process raw DICOM file from
373 echocardiography exam directly and output PE width prediction.

374 **Figure 3. Moving Window View Selection**

375 We used 48 continuous frames as window to iterate through raw video. In each
376 window, the view will be classified by majority voting from the prediction of all48
377 frames. For example, if most frames in the window were classified as view 2, the
378 window was classified as view 2. As the window iterate through raw video, the
379 window with highest confidence of classification will be preserved for further
380 analysis.

381

382 **Figure 4. Scatter plot of PE prediction from deep learning model compare with** 383 **human expert**

384 Figure 4A represent the scatter plot from internal validation, and Figure 4B from
385 external validation.

386

387 **Acknowledgements**

388 We appreciated the Biostatistics Center, Kaohsiung Chang Gung Memorial Hospital

389 for statistics work.

390

391 **Funding**

392 The study was supported by grants EDCHM109001 and EDCHP110001 from E-Da

393 Cancer Hospital and grants CMRPG8M0181 from the Chang Gung Medical

394 Foundation.

395

396 **Disclosures**

397 The authors declare that they have no competing interests.

398

399 **Reference**

- 400 1. Jung HO. Pericardial effusion and pericardiocentesis: role of
401 echocardiography. *Korean Circ J.* 2012;42:725-734.
402 doi:10.4070/kcj.2012.42.11.725
- 403 2. Ceriani E, Cogliati C. Update on bedside ultrasound diagnosis of pericardial
404 effusion. *Intern Emerg Med.* 2016;11:477-480.
405 doi:10.1007/s11739-015-1372-8
- 406 3. Vakamudi S, Ho N, Cremer PC. Pericardial Effusions: Causes, Diagnosis, and
407 Management. *Prog Cardiovasc Dis.* 2017;59:380-388.
408 doi:10.1016/j.pcad.2016.12.009
- 409 4. Malik SB, Chen N, Parker III RA, Hsu JY. Transthoracic echocardiography:
410 pitfalls and limitations as delineated at cardiac CT and MR imaging.
411 *Radiographics.* 2017;37:383-406.
- 412 5. Sampaio F, Ribeiros R, Galrinho A, Teixeira R, João I, Trábulo M, Quelhas I,
413 Cabral S, Ribeiro J, Mendes M, Morais J. Consensus document on
414 transthoracic echocardiography in Portugal. *Rev Port Cardiol (Engl Ed).*
415 2018;37:637-644. doi:10.1016/j.repc.2018.05.009

- 416 6. Sagristà-Sauleda J, Mercé AS, Soler-Soler J. Diagnosis and management of
417 pericardial effusion. *World journal of cardiology*. 2011;3(5):135.
418 doi:10.4330/wjc.v3.i5.135
- 419 7. de Siqueira VS, Borges MM, Furtado RG, Dourado CN, da Costa RM.
420 Artificial intelligence applied to support medical decisions for the automatic
421 analysis of echocardiogram images: A systematic review. *Artificial*
422 *Intelligence in Medicine*. 2021;120:102165.
423 doi:10.1016/j.artmed.2021.102165
- 424 8. Zhou J, Du M, Chang S, Chen Z. Artificial intelligence in echocardiography:
425 detection, functional evaluation, and disease diagnosis. *Cardiovascular*
426 *Ultrasound*. 2021;19:1-11. Doi: 10.1136/heartjnl-2021-319725
- 427 9. Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, Heidenreich
428 PA, Harrington RA, Liang DH, Ashley EA, Zou JY. Video-based AI for
429 beat-to-beat assessment of cardiac function. *Nature*. 2020;580:252-256. doi:
430 10.1038/s41586-020-2145-8
- 431 10. Huang M-S, Wang C-S, Chiang J-H, Liu P-Y, Tsai W-C. Automated
432 recognition of regional wall motion abnormalities through deep neural
433 network interpretation of transthoracic echocardiography. *Circulation*.
434 2020;142:1510-1520. doi: 10.1161/CIRCULATIONAHA.120.047530

- 435 11. Nayak A, Ouyang D, Ashley EA. A DEEP LEARNING ALGORITHM
436 ACCURATELY DETECTS PERICARDIAL EFFUSION ON
437 ECHOCARDIOGRAPHY. *Journal of the American College of Cardiology*.
438 2020;75(11_Supplement_1):1563-1563. doi: 10.1016/S0735-1097(20)32190-2
- 439 12. Zhang J, Gajjala S, Agrawal P, Tison GH, Hallock LA, Beussink-Nelson L,
440 Lassen MH, Fan E, Aras MA, Jordan C, Fleischmann KE, Melisko M, Qasim
441 A, Shah SJ, Bajcsy R, Deo RC. Fully automated echocardiogram
442 interpretation in clinical practice: feasibility and diagnostic accuracy.
443 *Circulation*. 2018;138:1623-1635. doi:
444 10.1161/CIRCULATIONAHA.118.034338.
- 445 13. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition.
446 Paper presented at: Proceedings of the IEEE conference on computer vision
447 and pattern recognition 2016. doi: 10.48550/arXiv.1512.03385
- 448 14. He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. Paper presented at:
449 Proceedings of the IEEE international conference on computer vision 2017.
450 doi: 10.48550/arXiv.1703.06870
- 451 15. Ghorbani A, Ouyang D, Abid A, He B, Chen JH, Harrington RA, Liang DH,
452 Ashley EA, Zou JY. Deep learning interpretation of echocardiograms. *NPJ*
453 *digital medicine*. 2020;3:1-10. doi: 10.1038/s41746-019-0216-8

- 454 16. Inglis R, King AJ, Gleave M, Adlam D. Pericardiocentesis in contemporary
455 practice. *Journal of Invasive Cardiology*. 2011;23:234–239. PMID: 21646649
- 456 17. Tsang TS, Seward JB, Barnes ME, Bailey KR, Sinak LJ, Urban LH, Hayes SN.
457 Outcomes of primary and secondary treatment of pericardial effusion in
458 patients with malignancy. *Mayo Clin Proc*. 2000;75:248-253. doi:
459 10.4065/75.3.248
- 460 18. Tsang TS, Enriquez-Sarano M, Freeman WK, Barnes ME, Sinak LJ, Gersh BJ,
461 Bailey KR, Seward JB. Consecutive 1127 therapeutic echocardiographically
462 guided pericardiocenteses: clinical profile, practice patterns, and outcomes
463 spanning 21 years. *Mayo Clin Proc*. 2002;77:429-436. doi: 10.4065/77.5.429.
- 464 19. Nguyen CT, Lee E, Luo H, Siegel RJ. Echocardiographic guidance for
465 diagnostic and therapeutic percutaneous procedures. *Cardiovascular diagnosis
466 and Therapy*. 2011;1:11-36. doi: 10.3978/j.issn.2223-3652.2011.09.02.
- 467 20. Maggolini S, Gentile G, Farina A, De Carlini CC, Lenatti L, Meles E, Achilli
468 F, Tempesta A, Brucato A, Imazio M. Safety, efficacy, and complications of
469 pericardiocentesis by real-time echo-monitored procedure. *The American
470 Journal of Cardiology*. 2016;117:1369-1374. doi:
471 10.1016/j.amjcard.2016.01.043

- 472 21. Ebert LC, Heimer J, Schweitzer W, Sieberth T, Leipner A, Thali M,
473 Ampanozi G. Automatic detection of hemorrhagic pericardial effusion on
474 PMCT using deep learning-a feasibility study. *Forensic Science, Medicine and*
475 *Pathology*. 2017;13:426-431. doi: 10.1007/s12024-017-9906-1
- 476 22. Draelos RL, Dov D, Mazurowski MA, Lo JY, Henao R, Rubin GD, Carin L.
477 Machine-learning-based multiple abnormality prediction with large-scale chest
478 computed tomography volumes. *Medical image analysis*. 2021;67:101857. doi:
479 10.1016/j.media.2020.101857
- 480 23. Khamis H, Zurakhov G, Azar V, Raz A, Friedman Z, Adam D. Automatic
481 apical view classification of echocardiograms using a discriminative learning
482 dictionary. *Medical image analysis*. 2017;36:15-21. doi:
483 10.1016/j.media.2016.10.007
- 484 24. Sengupta PP, Huang YM, Bansal M, Ashrafi A, Fisher M, Shameer K, Gall W,
485 Dudley JT. Cognitive machine-learning algorithm for cardiac imaging: a pilot
486 study for differentiating constrictive pericarditis from restrictive
487 cardiomyopathy. *Circulation: Cardiovascular Imaging*. 2016;9:e004330. doi:
488 10.1161/CIRCIMAGING.115.004330

- 489 25. Gao X, Li W, Loomes M, Wang L. A fused deep learning architecture for
490 viewpoint classification of echocardiography. *Information Fusion*.
491 2017;36:103-113. doi: 10.1016/j.inffus.2016.11.007
- 492 26. Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view
493 classification of echocardiograms using deep learning. *NPJ digital medicine*.
494 2018;1:1-8. doi: 10.1038/s41746-017-0013-1
- 495 27. Karia DH, Xing Y-Q, Kuvin JT, Nesser HJ, Pandian NG. Recent role of
496 imaging in the diagnosis of pericardial disease. *Current Cardiology Reports*.
497 2002;4:33-40. doi: 10.1007/s11886-002-0124-3. PMID: 11743920
- 498 28. Prakash AM, Sun Y, Chiaramida SA, Wu J, Lucariello RJ. Quantitative
499 assessment of pericardial effusion volume by two-dimensional
500 echocardiography. *Journal of the American Society of Echocardiography*.
501 2003;16:147-153. doi: 10.1067/mje.2003.35
- 502 29. Pérez-Casares A, Cesar S, Brunet-Garcia L, Sanchez-de-Toledo J.
503 Echocardiographic evaluation of pericardial effusion and cardiac tamponade.
504 *Frontiers in pediatrics*. 2017;5:79. doi: 10.3389/fped.2017.00079

505 Table 1. Demographics, Basic Characteristics, and Clinical Findings of the patients

Variables	Training set	Internal Validation set	External Validation set
Number of patients	582	155	258
Age(y)	67.4±15.4	59.8±19.2	66.4±16.1
Gender	Male 49.5%	Male 53.2%	Male 58.1%
Height, cm	157.9±9.2	162.1±8.4	161.6±8.4
Weight, kg	53.7±35.6	65.9±16.8	61.8±13.3
BMI, kg/m ²	23.3±5.2	24.4±1.5	23.6±4.5
Mode EF %	61.1±13.8	64.3±7.1	61±13.9
Mode EF<50 (%)	12.8%	8.4%	19.3%
patients with PE	582(100%)	72(46.5%)	163(63.2)

506

507

508 Table 2. Dice coefficient of image segmentation

	Dice Coefficient					
	PE	RV	LV	RA	LA	Average
PLAX	0.72	0.86	0.85		0.84	0.82
PSAX	0.69	0.59	0.85			0.71
A4C	0.58	0.81	0.86	0.82	0.83	0.78
SC	0.56	0.66	0.71	0.70	0.72	0.67

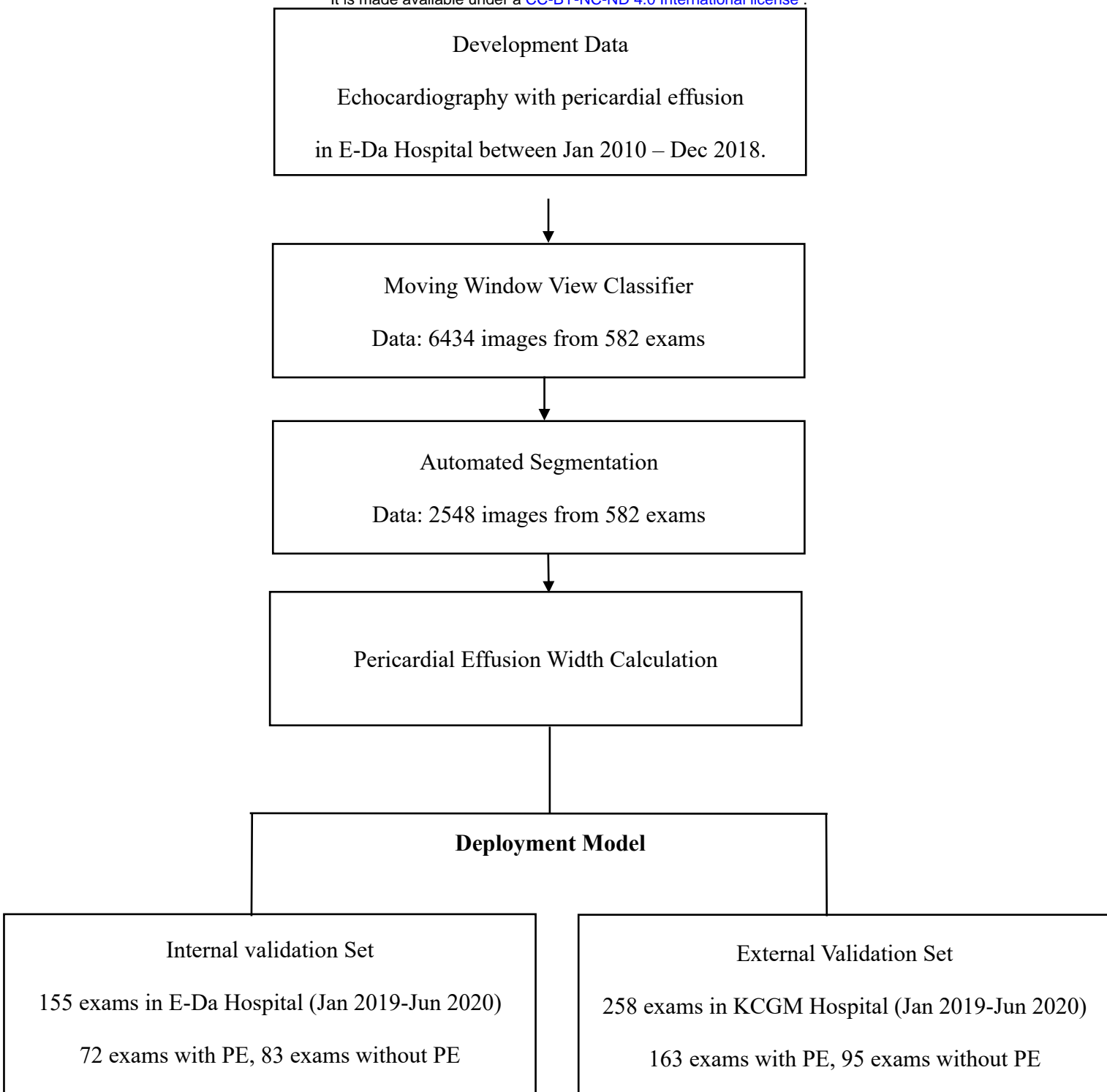
509

510

511 Table 3.

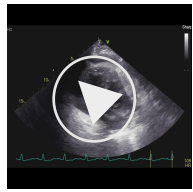
	number before view selection	number after moving window selection	Mean Absolute Error (cm)	ICC	R2
Internal validation					
PLAX	n=155	n=146	0.28	0.910 (0.876-0.935)	0.700
PSAX	n=155	n=155	0.46	0.802 (0.728-0.856)	0.469
A4C	n=155	n=155	0.32	0.907 (0.871-0.932)	0.754
SC	n=155	n=124	0.40	0.865 (0.808-0.905)	0.590
External validation					
PLAX	222	n=127	0.32	0.807 (0.726-0.864)	0.457
PSAX	252	n=138	0.44	0.714 (0.600-0.796)	0.337
A4C	212	n=100	0.11	0.897 (0.846-0.931)	0.662

512



medRxiv preprint doi: <https://doi.org/10.1101/2022.08.13.22278732>; this version posted August 16, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-NC-ND 4.0 International license](#) .

Input



(A) Model 1: ResNet 50

view selection
image

7×7 conv, 64, /2

pool, /2

3×3 conv, 64

3×3 conv, 64

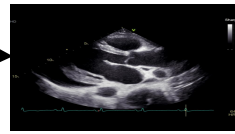
3×3 conv, 512

3×3 conv, 512

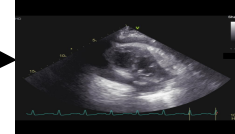
avg pool

fc 1000

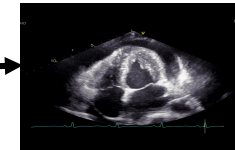
PLAX



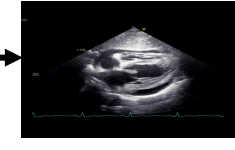
PSAX



A4C



SC

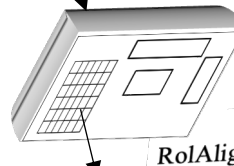
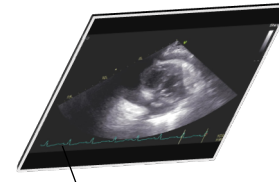


Low confidence videos exclude

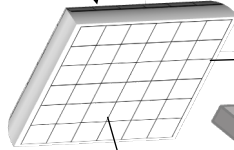


(B) Model 2: Mask R-CNN

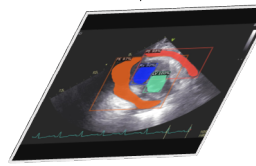
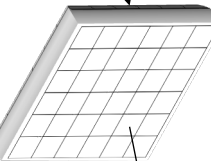
Automated segmentation



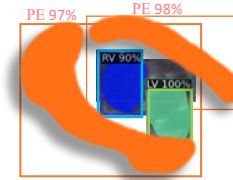
RoIAlign



class box



(C) Pericardial effusion prediction



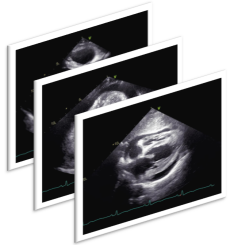
PE grading

Non-ML
calculate PE
width from
segmentation

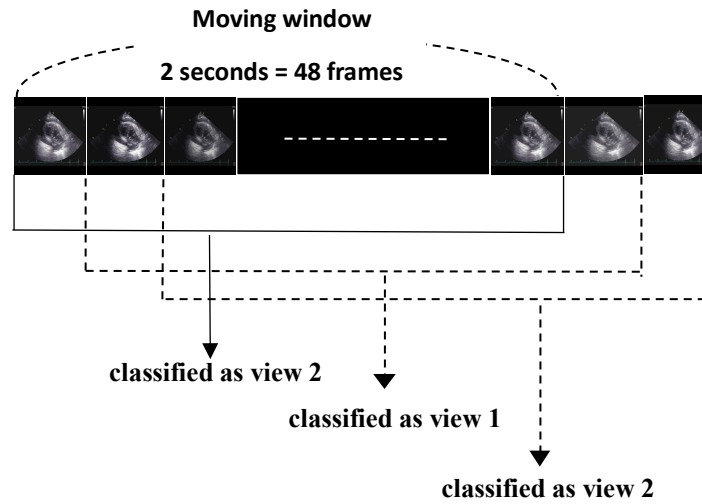


Raw data

Continuous frames
from DICOM file



➔ **Model 1 ResNet 50 :**



Majority voting

most windows in a video classified as view 2

⇒ **the video is view 2**

