Automated, Vendor-agnostic Measurement of Myocardial Tissue Velocities in Echocardiography

Waqar Ali* ⓘ Shana Stämpfli† Thilo Stadelmann*‡ ⓘ Schlomo Aschkenasy† Ahmed Abdulkadir* ⓘ

{aliw,stdm,abdk}@zhaw.ch, schlomo@medirapp.ai

* Centre for Artificial Intelligence, Zurich University of Applied Sciences, Winterthur, Switzerland
† MediRapp AG, Zürich, Switzerland
‡Fellow, ECLT (European Centre for Living Technology, Venice, Italy) and Senior Member, IEEE

Abstract—In this study, we present a vendor-agnostic, deep learning-based system for the automated analysis of transthoracic pulsed-wave tissue Doppler imaging (TDI), which decouples image acquisition from interpretation and enables centralized, fleet-wide analysis across devices. The model ingests standard TDI from heterogeneous ultrasound systems and automatically extracts key diagnostic markerspeak systolic velocity (S'), early diastolic velocity (e'), and late diastolic/atrial contraction velocity (a') using a single, unified pipeline. Conceptually, this harmonizes measurements across vendors and sites, improving consistency, comparability, and longitudinal tracking without device-specific calibration or tooling. Procedurally, a central inference service supports asynchronous batch processing and human-in-the-loop review, thereby shifting analysis off-console, allowing ultrasound scanners to remain fully available for acquisition. In our clinical dataset, which spans two ultrasound vendors and diverse cardiac cycles, the system correctly identified more than 93% of tissue-velocity landmarks. In 50% of studies, all automated detections matched expert annotations, eliminating the need for manual edits. This approach streamlines offline TDI analysis, accelerates turnaround, and supports scalable, standardized cardiac assessments.

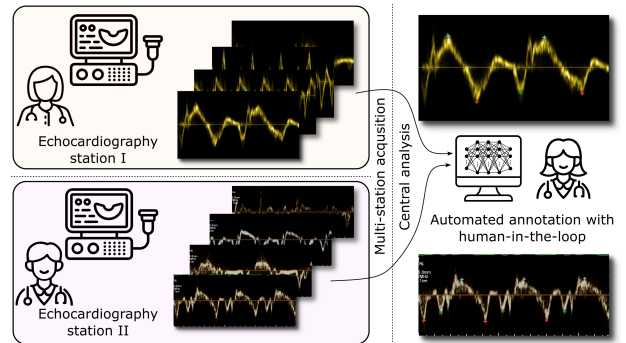*Index Terms*—echocardiography, tissue Doppler imaging, deep learning, clinical application

Fig. 1: Vendor-neutral automated analysis of peak velocities in echocardiographic tissue Doppler imaging in which the acquisitions are performed with two different stations (left) and subsequently analyzed in a central location without needing access to an acquisition device (right).

I. INTRODUCTION

As part of an echocardiographic assessment, pulsed-wave tissue Doppler imaging (TDI) [1], [2] is employed to measure myocardial tissue velocities, which are instrumental in diagnosing conditions such as diastolic dysfunction and other adverse heart conditions [3], [4]. Key diagnostically relevant metrics include the peak systolic velocity (S'), early diastolic velocity (e'), and late diastolic or atrial contraction velocity (a'), which are critical for evaluating both systolic and diastolic heart function. According to international echocardiographic guidelines, these velocities are measured at the mitral annulus using TDI [5]–[7]. They are quantified through analysis of the spectral Doppler waveform displayed on ultrasound monitors. As the annotation of a single image typically requires up to a minute, and patients often undergo multiple TDI examinations during a single visit, this manual process detracts from the primary diagnostic workflow.

Current manual annotation tools and proprietary automated solutions require echocardiography machines to serve both acquisition and annotation tasks, limiting their availability for additional patients. Vendor-agnostic automated annotation enhances workflow efficiency by pre-processing TDI data, enabling cardiologists to begin interpretation immediately without the cold start associated with reviewing unannotated images acquired by sonographers. This approach preserves scanner availability, accelerates measurement collection, and reduces the overall time required for diagnostic assessment in high-demand clinical settings.

Related research has explored deep learning ap-

proaches for automated echocardiographic analysis. [8] demonstrated that deep learning algorithms, trained on labeled 2D videos and Doppler images, can accurately measure cardiac structural and Doppler parameters, achieving high agreement with human experts while substantially reducing analysis time. [9] extended this approach to large, multi-center datasets, showing that fully automated workflows could classify, segment, and annotate 2D and Doppler echocardiograms with high accuracy, consistent measurements, and reduced variability compared with human readers. While these studies show the potential of automated image analysis in cardiology, we focus on the evaluation of guideline-confirming, verifiable measurements. This narrower field of research has been largely restricted to single-device analyses [10] or demonstrates limited accuracy [11]–[13]. In contrast, this study presents a manufacturer-agnostic system that automatically identifies S', e', and a' velocities on TDI images following international guidelines [5]–[7], streamlining the diagnostic process while freeing acquisition devices from prolonged non-productive use. By enabling clinicians to operate across multiple devices concurrently, the system has the potential to substantially improve clinical throughput and standardize measurements without compromising diagnostic precision. Fig. 1 illustrates the principle of multi-station echocardiographic acquisition combined with centralized, automated annotation within a human-in-the-loop workflow.

II. METHODS

A. Clinical data and expert annotations

The study included a retrospectively acquired dataset comprising of 664 ultrasound Doppler tissue images recorded with two different devices during routine clinical visits. After the removal of two incomplete images and 90 (near) duplicates, 572 TDI images were included in this study. Five hundred and thirty-nine of those images were acquired on a device manufactured by General Electrics (GE) and 33 on a device manufactured by Philips with varying image sizes (approximately 800x450 pixels). All participants of the study gave written consent that the anonymized data can be used for research and to train AI models. The dataset includes a diverse range of patients with different cardiac conditions. This ensures the representation of regular and irregular heart cycles. Each image was annotated by the same expert cardiologist (S.A.) using custom labeling software, designed to mark S', e', and a' peak velocities directly on the echocardiographic images. The offset and scaling to convert pixel coordinates to physical coordinates of time in seconds and velocity in $\mathrm{cm/s}$ was part of the vendor-specific data extraction step.

B. Deep learning model and training

We solve the localization task as a segmentation task. The segmentation masks are generated by marking the labeled points with a circular mask with a diameter of 31 pixels, one mask for each type of peak velocity. The background mask is then created as the inverse of the union of the three peak tissue velocity masks. Stacking all masks results in a four-channel map that serves as ground truth. We employed a U-Net with padded convolutions but otherwise identical architecture as the original implementation [14]. The U-Net is recognized for its ability to perform accurate pixel-level segmentation and capture both global context and fine-grained details. We optimize the Dice loss [15] because it is more robust than the cross-entropy loss against class imbalances [16]. Inversely to the generation of the mask of the ground truth, the predicted coordinates were computed from the output of the U-Net by finding their centers of mass of predicted components. The architecture can process input images of arbitrary size. To obtain correct physical units in $\mathrm{cm/s}$, the affine pixel-to-world transformation obtained from the original DICOM (Digital Imaging and Communications in Medicine) data is used.

The dataset was randomly split into training and testing sets, with the test set comprising 80 images (8 Philips images and 72 GE images) and the training set containing 492 images (25 Philips images and 467 GE images). This ensures that the model is evaluated on unseen data. The training was performed using a batch size of 8 for 75 epochs. The trained model was validated using a test set containing Doppler ultrasound images, and segmentation masks were generated for each class. Additionally, we conducted a qualitative assessment by visually comparing the predicted segmentation masks with overlaid ground truth markers, ensuring a comprehensive evaluation of the model's performance.

To generalize to data from multiple vendors, we incorporate color- and intensity-based augmentation including permutation of color channels, grayscale transformation, and variation of intensity mapping that modifies the images' brightness, contrast, and color/grayscale appearance, simulating variations in Doppler image appearance of different display styles. The color jitter augmentation ($A_{\mathrm{jit.}}$) randomly jitters the brightness factor, contrast factor, saturation, and hue prducing a wide range of appearances. The grayscale or permutation augmentation ($A_{\mathrm{p.g.}}$) randomly shuffles color channels or converts images to grayscale with equal probability. The combination of the above ($A_{\mathrm{comb.}}$) combines $A_{\mathrm{jit.}}$ and $A_{\mathrm{p.g.}}$ by selecting either one augmentation with equal probability and thus produces the largest variety in the input.

C. Evaluation of the detection rate and accuracy

Model performance was evaluated at two hierarchical levels. At the waveform level, each individual TDI spectral waveform was assessed, focusing on the detection of S', e', and a' markers, typically three or more per image. At the patient level, performance was assessed across all waveforms within a given patients study, providing an aggregate measure of the systems accuracy for complete examinations. We define a marker placed by the domain expert as *detected* if there is a predicted marker of the same type within $\pm 0.2\,\mathrm{s}$ and within $\pm 2.5\,\mathrm{cm/s}$. From this rule, we define following quantitative performance metrics:

The waveform-level true positive rate (TPR_{waveform}) is the fraction of manually annotated peak velocity markers of a given class that have at least one match.

The waveform-level false positive rate (FPR_{waveform}) is the number of predictions that do not have a matching ground truth marker divided by the number of ground truth markers.

The patient-level true positive rate (TPR_{patient}) is the fraction of images of the test set in which the TPR_{waveform} is 1.

Patient-level no false negatives ($noFN_{\mathrm{patient}}$) is the fraction of images of the test set that have no false negatives.

The patient-level no false positives ($noFP_{\mathrm{patient}}$) is the fraction of images of the test set in which FPR_{waveform} is zero.

Patient-level no errors (noE_{patient}) is the fraction of images of the test set in which TPR_{waveform} is 1 and no false positives are present.

III. RESULTS

The baseline network detected more than 90 % of the markers (Table I). This result improved by about two percent with the color jitter augmentation. However, no augmentation method dominated all others. Training the network without augmentation was in none of the metrics optimal.

On the patient level, the U-Net trained with color jitter predicted 50% of the TDI without errors (Table I). For individual peak tissue velocity phases, the TPR_{patient} was between 0.825 and 0.925.

On the waveform level, color jitter augmentation had the lowest FPR_{waveform}. Clinically, reducing the number of false positives is crucial to reducing inaccurate velocity estimations, reinforcing that color jitter augmentation provides the best alternative for this particular task.

On average, the peak tissue velocities were predicted with a bias of $-0.03\,\mathrm{cm/s}$, $-0.07\,\mathrm{cm/s}$, $0.18\,\mathrm{cm/s}$ and a standard deviation of $0.60\,\mathrm{cm/s}$, $0.60\,\mathrm{cm/s}$, and $0.47\,\mathrm{cm/s}$ for (e'), (a'), and (S'), respectively (Fig. 2).

TABLE I: Detection performance on waveform level obtained with different augmentation methods (top four rows) and the patient-level performance using intensity augmentation (bottom row).

		Peak velocity measure			
Augm.	**Metric**	**S'**	**e'**	**a'**	**All**
none	TPR_{waveform}	0.924	0.899	0.913	0.914
	FPR_{waveform}	0.138	0.135	0.134	0.136
$A_{\mathrm{jit.}}$	TPR_{waveform}	**0.960**	0.919	0.919	**0.933**
	FPR_{waveform}	**0.122**	**0.129**	**0.119**	**0.124**
$A_{\mathrm{p.g.}}$	TPR_{waveform}	0.950	0.889	**0.929**	0.923
	FPR_{waveform}	0.149	0.133	0.149	0.136
$A_{\mathrm{comb.}}$	TPR_{waveform}	0.925	**0.924**	0.918	0.922
	FPR_{waveform}	0.132	0.133	0.151	0.131
$A_{\mathrm{jit.}}$	TPR_{patient}	0.925	0.875	0.825	0.725
	$noFN_{\mathrm{patient}}$	0.937	0.875	0.850	0.725
	$noFP_{\mathrm{patient}}$	0.800	0.787	0.762	0.562
	noE_{patient}	0.750	0.725	0.637	0.500

none: no augmentation; $A_{\mathrm{jit.}}$: intensity augmentation; $A_{\mathrm{p.g.}}$: color channel permutation or grayscale augmentation; $A_{\mathrm{comb.}}$: combination of augmentations. The table reports the performance on the waveform level, specifically true positive rate (TPR_{waveform}) and false positive (FPR_{waveform}) for S', e', a' waves, and overall. The best values in each column are highlighted in bold.
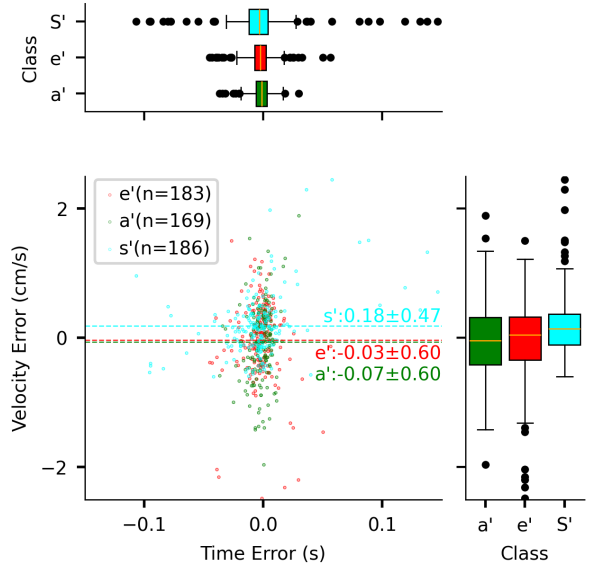

Fig. 2: Deviation of the predicted myocardial peak tissue velocity and time to the corresponding matching ground truth velocity and time. The dashed lines indicate the mean deviation of each class. The boxplots show the median (yellow line), a box with interquartile distance (IQD), whiskers at $\pm 2 \times$IQD or the extreme value. Black dots beyond the whiskers indicate outlier values further away than $\pm 2 \times$IQD.
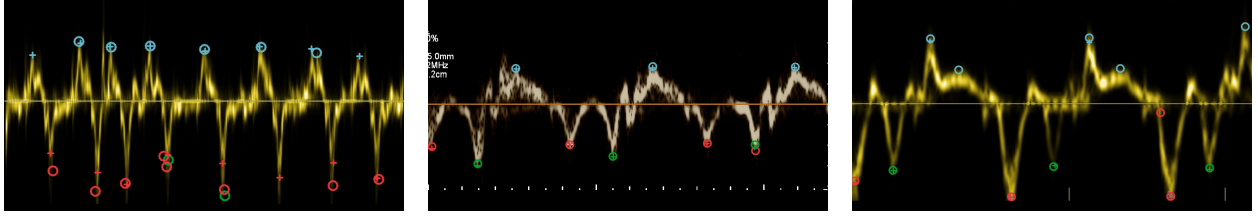
Fig. 3: Examples of incorrect patient-level predictions. The ground truths (+) and predictions (o) of tissue velocity peaks S (cyan), e' (red), and a' (green) are overlaid on the input image. Left: Unusual signal in which no a' waves are preset. The model confuses a' and e' and misses two S' waves (false . Centre: The last a' wave is marked as a' and e' (false positive). Right: There are two S waves marked in two of the waveforms and two e' waves marked in the last waveform.

IV. DISCUSSION

Deep learningbased detection of peak tissue velocities in TDI has been reported previously [10], [17]. While direct comparison with these studies is not feasible due to differences in test datasets, our approach measures the same clinically relevant physical quantities cm/s, ensuring that the results remain directly meaningful for practical diagnostic applications.

For matched points, we found measurement errors ranging from $\pm 0.47\,cm/s$ to $\pm 0.60\,cm/s$ while [10] found errors ranging from $\pm 0.26\,cm/s$ to $\pm 1.25\,cm/s$ between the consensus of two experts and a third expert or between the consensus of two experts and the machine [10]. Overall, the error magnitudes observed in our study were comparable to those reported in prior work, though they exceeded those observed in previous inter-observer comparisons between human experts.

At the patient level, all peak tissue velocity points for each wave type were accurately detected without any false positives in 50% of the cases. This implies that in the clinical setting, we expect that in half the cases no intervention is necessary to obtain quantitative values that are not inferior to those from a domain expert but with less effort.

In the qualitative evaluation of failure cases (Fig. 3), we found that not all errors are critical or true errors. For example, some unmarked points at the edge of the image were likely missed in the ground truth. Their inclusion would not affect the extracted patient-level diagnostic value. Confusions between e' and a' are considered critical. These cases, however, could be flagged with algorithmic plausibility checks requiring a strict order of $S' \rightarrow e' \rightarrow a'$. Also considered critical are completely misplaced markers. This happens typically in unusual (e.g. low-quality) signals and can also be flagged algorithmically whenever the variability of the velocity within a class is too large.

Despite the possibility of flagging failures with an algorithm, the application in its current state requires a human in the loop to verify, and, if needed, correct the proposal. Human verification is also needed because the accuracy is lower than that of human domain experts.

V. CONCLUSION

Our study presents a vendor-agnostic deep learning system for automated detection of peak myocardial tissue velocities in TDI, designed to streamline echocardiographic workflows without compromising diagnostic accuracy. The model operates robustly across heterogeneous imaging systems and arbitrary image sizes without the need for vendor-specific retraining, enabled by augmentation-heavy training and standardized preprocessing.

Using a representative clinical dataset from two echocardiographic platforms prevalent in Switzerland, the system correctly identified over 93% of tissue-velocity landmarks. In half of the studies, all detected peaks met predefined quality thresholds, substantially reducing the need for manual corrections and accelerating analysis throughput. Errors such as confusion between the e' and a' TDI waves were rare and reliably captured by automated post-processing checks, ensuring that uncertain outputs are flagged for expert review.

Integrated within a human-in-the-loop workflow, the system enables continuous scanner availability, supports parallel operation of multiple workstations, and promotes consistent, vendor-neutral interpretation of TDI data. These features contribute to improved operational efficiency and standardization in echocardiographic assessment.

Future work will focus on extending generalization to additional ultrasound vendors, integrating quantitative quality metrics for automated confidence estimation, and validating the system across broader and more diverse patient populations. Ultimately, this approach aims to advance reproducibility, scalability, and accessibility of echocardiographic tissue velocity analysis in routine clinical practice.

REFERENCES

[1] D. Vinereanu, A. Khokhar, and A. G. Fraser, "Reproducibility of Pulsed Wave Tissue Doppler Echocardiography," *Journal of the American Society of Echocardiography*, vol. 12, no. 6, pp. 492–499, Jun. 1999.

[2] U. G. Sisto, D. Orso, D. Maione, F. Venturelli, and A. De Luca, "Tissue Doppler Imaging in Acute and Critical Care: Enhancing Diagnostic Precision," *Medicina*, vol. 61, no. 6, p. 1051, Jun. 2025.

[3] L. Galiuto, G. Ignone, and A. N. DeMaria, "Contraction and Relaxation Velocities of the Normal Left Ventricle Using Pulsed-Wave Tissue Doppler Echocardiography," *The American Journal of Cardiology*, vol. 81, no. 5, pp. 609–614, Mar. 1998.

[4] K. K. Kadappu and L. Thomas, "Tissue Doppler Imaging in Echocardiography: Value and Limitations," *Heart, Lung and Circulation*, vol. 24, no. 3, pp. 224–233, Mar. 2015.

[5] W. A. Zoghbi, J. B. Chambers, J. G. Dumesnil, E. Foster, J. S. Gottdiener, P. A. Grayburn, B. K. Khandheria, R. A. Levine, G. R. Marx, F. A. Miller, S. Nakatani, M. A. Quiñones, H. Rakowski, L. L. Rodriguez, M. Swaminathan, A. D. Waggoner, N. J. Weissman, and M. Zabalgoitia, "Recommendations for Evaluation of Prosthetic Valves With Echocardiography and Doppler Ultrasound," *Journal of the American Society of Echocardiography*, vol. 22, no. 9, pp. 975–1014, Sep. 2009.

[6] S. F. Nagueh, O. A. Smiseth, C. P. Appleton, B. F. Byrd, H. Dokainish, T. Edvardsen, F. A. Flachskampf, T. C. Gillebert, A. L. Klein, P. Lancellotti, P. Marino, J. K. Oh, B. A. Popescu, and A. D. Waggoner, "Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging," *Journal of the American Society of Echocardiography*, vol. 29, no. 4, pp. 277–314, Apr. 2016.

[7] G. Zamzmi, L.-Y. Hsu, W. Li, V. Sachdev, and S. Antani, "Fully automated spectral envelope and peak velocity detection from Doppler echocardiography images," in *Medical Imaging Computer-Aided Diagnosis*, vol. 11314, Mar. 2020, pp. 1053–1064.

[8] G. Peng, R. Ling, X. Liu, Q. Liu, X. Zhong, Y. Sheng, Y. Zheng, S. Luo, Y. Yang, X. Lin, K. Tang, J. Zheng, L. Chen, D. Ni, J. Xu, Y. Liu, and W. Xue, "Formal validation of a deep learning-based automated interpretation system for cardiac structure and function in adult echocardiography," *Quantitative Imaging in Medicine and Surgery*, vol. 15, no. 4, pp. 3093–3110, Apr. 2025.

[9] J. Tromp, P. J. Seekings, C.-L. Hung, M. B. Iversen, M. J. Frost, W. Ouwerkerk, Z. Jiang, F. Eisenhaber, R. S. M. Goh, H. Zhao, W. Huang, L.-H. Ling, D. Sim, P. Cozzone, A. M. Richards, H. K. Lee, S. D. Solomon, C. S. P. Lam, and J. A. Ezekowitz, "Automated interpretation of systolic and diastolic function on the echocardiogram: a multicohort study," *The Lancet Digital Health*, vol. 4, no. 1, pp. e46–e54, Jan. 2022.

[10] E. S. Lane, J. Jevsikov, M. J. Shun-shin, N. Dhutia, N. Matoorian, G. D. Cole, D. P. Francis, and M. Zolgharni, "Automated multi-beat tissue Doppler echocardiography analysis using deep neural networks," *Medical & Biological Engineering & Computing*, vol. 61, no. 5, pp. 911–926, May 2023.

[11] N. M. Dhutia, M. Zolgharni, M. Mielewczik, M. Negoita, S. Sacchi, K. Manoharan, D. P. Francis, and G. D. Cole, "Open-source, vendor-independent, automated multi-beat tissue Doppler echocardiography analysis," *The International Journal of Cardiovascular Imaging*, vol. 33, no. 8, pp. 1135–1148, Aug. 2017.

[12] M. Y. Elwazir, Z. Akkus, D. Oguz, Z. Ye, and J. K. Oh, "Fully Automated Mitral Inflow Doppler Analysis Using Deep Learning," in *IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, Oct. 2020, pp. 691–696.

[13] A. Taebi, R. H. Sandler, B. Kakavand, and H. A. Mansy, "Estimating Peak Velocity Profiles from Doppler Echocardiography using Digital Image Processing," in *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Dec. 2018, pp. 1–4.

[14] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention MICCAI*, Cham, 2015, pp. 234–241.

[15] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, Oct. 2016, pp. 565–571.

[16] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021.

[17] J. Jevsikov, T. Ng, E. S. Lane, E. Alajrami, P. Naidoo, P. Fernandes, J. S. Sehmi, M. Alzetani, C. D. Demetrescu, N. Azarmehr, N. D. Serej, C. C. Stowell, M. J. Shun-Shin, D. P. Francis, and M. Zolgharni, "Automated mitral inflow Doppler peak velocity measurement using deep learning," *Computers in Biology and Medicine*, vol. 171, p. 108192, Mar. 2024.