Automatic Detection of Aortic Valve Events Using Deep Neural Networks on Cardiac Signals From Epicardially Placed Accelerometer

Ali Wajdan , Tollef Struksnes Jahren , Manuel Villegas-Martinez , Faraz H Khan,
Per Steinar Halvorsen, Hans Henrik Odland , Ole Jakob Elle, Anne H Schistad Solberg ,
and Espen W Remme

Abstract—Background: Miniaturized accelerometers incorporated in pacing leads attached to the myocardium, are used to monitor cardiac function. For this purpose functional indices must be extracted from the acceleration signal. A method that automatically detects the time of aortic valve opening (AVO) and aortic valve closure (AVC) will be helpful for such extraction. We tested if deep learning can be used to detect these valve events from epicardially attached accelerometers, using high fidelity pressure measurements to establish ground truth for these valve events. Method: A deep neural network consisting of a CNN, an RNN, and a multi-head attention module was trained and tested on 130 recordings from 19 canines and 159 recordings from 27 porcines covering different interventions. Due to limited data, nested cross-validation was used to assess the accuracy of the method. Result: The correct detection rates were 98.9% and 97.1% for AVO and AVC in canines and 98.2% and 96.7% in porcines when defining a correct detection as a prediction closer than 40 ms to the ground truth. The incorrect detection rates were 0.7% and 2.3% for AVO and AVC in canines and 1.1% and 2.3% in porcines. The mean absolute error between correct detections and their ground truth was 8.4 ms and 7.2 ms for AVO and AVC in canines, and 8.9 ms and 10.1 ms in porcines. Conclusion: Deep neural networks can be used on signals from epicardially attached accelerometers for robust and accurate detection of the opening and closing of the aortic valve.

Index Terms—Accelerometer, aortic valve opening, aortic valve closure, deep learning, heart function.

Ali Wajdan and Ole Jakob Elle are with the Intervention Centre, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway, and also with the Department of Informatics, University of Oslo, 0372 Oslo, Norway (e-mail: ali.wajdan@outlook.com; oelle@ous-hf.no).

Tollef Struksnes Jahren and Anne H Schistad Solberg are with the Department of Informatics, University of Oslo, 0373 Oslo, Norway (e-mail: tollefsj@ifi.uio.no; anne@ifi.uio.no).

Manuel Villegas-Martinez and Per Steinar Halvorsen are with the Intervention Centre, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway, and also with the Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0372 Oslo, Norway (e-mail: manuel.villegas0@gmail.com; sthalvor@ous-hf.no).

Hans Henrik Odland is with the Department of Cardiology, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway (e-mail: hanshenrikodland@gmail.com).

Faraz H Khan is with the Institute for Surgical Research, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway, and also with the Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, 0372 Oslo, Norway (e-mail: faraz.hkh@gmail.com).

Espen W Remme is with the Intervention Centre, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway, and also with the Institute for Surgical Research, Oslo University Hospital, Rikshospitalet, 0372 Oslo, Norway (e-mail: espen.remme@medisin.uio.no).

Digital Object Identifier 10.1109/JBHI.2022.3181148

I. INTRODUCTION

IN RECENT years, accelerometers have been miniaturized enough to be incorporated in devices such as pacing electrodes attached to the heart [1], [2]. As the function of the heart is directly linked to motion, accelerometers attached on the heart can be used for monitoring changes in heart function [3]. While the acceleration signal has a complex waveform with multiple oscillations during the cardiac cycle, integration once to velocity and twice to displacement provides smoother waveforms similar to velocity, displacement and strain waveforms obtained by echocardiography. Attaching the accelerometer directly to the heart, therefore allows extraction of functional information at a level comparable to cardiac imaging. In contrast, most previous studies on cardiac use of accelerometers have been focused on non-invasive measurements of the vibrations on the skin transmitted through layers of tissue by the beating heart, so called, seismocardiography (SCG) [4]. While the myocardial displacement of typically 1 cm produces accelerations of about 1 g (g = 9.81 m/s^2), the skin vibrations have much lower amplitude and produce an acceleration typically measured in milli-g [5].

Invasive cardiac accelerometers are a relatively new technology and currently the only commercial production for clinical use is for cardiac resynchronization therapy where the sensor is attached in the right atrium and ventricle [1], [2]. Our group has proposed to incorporate such a sensor in the temporary pacemaker leads that are routinely attached to the epicardium during cardiac surgery [3]. This setup provides a novel method to monitor cardiac function, without any additional surgical procedure, during and after a cardiac surgery. The temporary pacemaker leads remain attached during the post-operative days, which is a critical phase where the patient needs to be continuously

monitored. The incorporated 3-axes accelerometer offers direct and continuous measurements of the heart motion in this period, and may have added value as for example motion abnormalities is often the first sign of dysfunction occurring prior to other signs such as changes in ECG [3], [6], [7], while imaging such as echocardiography cannot be used for continuous monitoring.

The acceleration signal is of limited suitability for direct clinical interpretation for functional status of the heart due to its complex waveform. Therefore, the signal must be processed to extract simple indices that reflect cardiac function. This process typically involves identification of specified time points in each cardiac cycle where the indices are extracted. In previous studies using epicardially attached accelerometers, early systolic velocity and displacement extracted from accelerometer signals were used for detecting ischemic events with high accuracy [7]. However, a reference time-point in the cardiac cycle is a prerequisite to enable the extraction of such functional indices. In the examples above, R-peaks from simultaneously recorded ECG were used to mark the point of start systole in the recorded acceleration signals.

While start systole is a relatively easy time-point to extract due to its coincidence with the R-peak, other time points in the cardiac cycle can be of high clinical interest to detect for calculation of other functional indices. Several studies have shown that post-systolic motion indices are helpful for assessment of functional abnormalities [8]–[10] which requires the detection of aortic valve closing (AVC). Additionally, the time-point of aortic valve opening (AVO) can be of interest which marks the end of the isovolumic contraction phase. For example, prolongation of the pre-ejection phase, from electrical activation to AVO, is a sign of reduced contractile function [11], [12] and it can be quantified if the time point of AVO is known.

Opening and closing of the valves cause transitions in myocardial motion, and distinct oscillations in the accelerometer signals typically occur at these time-points. If there are characteristic features in the signal that occur at these events, algorithms that recognize these features may be developed and used to detect the associated valve events. Several recent studies on automatic detection of valve events by prominent features in SCG signals have been performed using different approaches such as temporal enveloping [13]–[15], continuous wavelet transform [16], machine learning [17], [18], or signal processing combinations [19]. We have also proposed a signal processing method to detect mitral and aortic valve events on measurements from epicardially attached accelerometers [20]. Typically, these SCG studies have used expert opinion based annotation of recognizable features in the acceleration signals and the methods have been tested mainly in normal individuals and in few types of conditions with abnormal cardiac motion. For an automatic detection algorithm to be useful in clinical practice, there must be limited variations in the features between patients and when cardiac function changes due to abnormalities or medical interventions. Furthermore, such algorithms may also require a good ECG signal where R-peak and possibly T-wave detections are initially used to generate limited search windows for the desired features. However, over the recent years we have performed several animal studies collecting data from epicardially attached accelerometers under a variety of cardiac
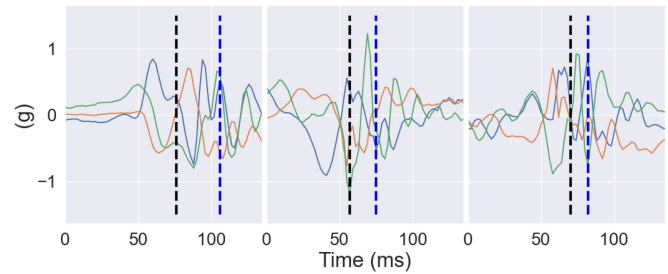
Fig. 1. Example cases showing the lack of common features for determining aortic valve opening in the early systolic complex. The colored waveforms from the 3-axis accelerometer represent measurements from the three different heart coordinates: Blue: longitudinal direction, Orange: circumferential direction, Green: radial direction. Black vertical dashed lines: time point of ECG R-peak, Blue vertical dashed lines: time point of aortic valve opening as defined by high fidelity pressure measurements.

interventions that represent the changes in heart function we expect to see in patients. From these data we have observed a large change in the features so we have not found repeatable features appearing consistently at the valve events which could be used for annotation or detection of the events. Fig. 1 illustrates recordings from three cases, demonstrating the challenge in defining common features.

Deep learning is an alternative method to expert opinion based feature extraction as this methodology may go deeper in the level of abstraction to a point beyond where humans can define the features, which means that features that may not be visible to the human brain are "visible" to the computer. Some recent articles have adopted deep learning for detection of valve events from echocardiographic images [21], [22]. Furthermore, in the case of ECG, deep learning has been extensively used to classify and segment the signals, for example detection of R-peak and segmentation of QRS waves [23], [24]. Machine and deep learning have also been used to quantify cardiac function from non-invasive wearable accelerometer and gyro signals [25], [26].

In this study, as a proof of concept, we have developed a deep neural network for automatic detection of aortic valve opening and closing from epicardially attached accelerometer signals and tested its performance under numerous interventions with varying cardiac function. The study has been performed using data recorded in previous canine and porcine experiments carried out by our group where simultaneously acquired left ventricular (LV) pressure was used as ground truth for annotating the true valve events. Furthermore, the derived method did not depend on ECG, which avoids problems with missing or bad ECG signal.

II. METHODOLOGY

A. Data Acquisition

The study was performed on data taken from previous experiments performed at Oslo University Hospital. All protocols were approved by the Norwegian Food Safety Authority (FOTS), and all experiments were carried out in accordance with Norwegian regulations concerning the use of animals in experiments. Data comprised recordings from 19 mongrel canines [FOTS ID: 17644] and 27 Norwegian Landrace porcines [FOTS
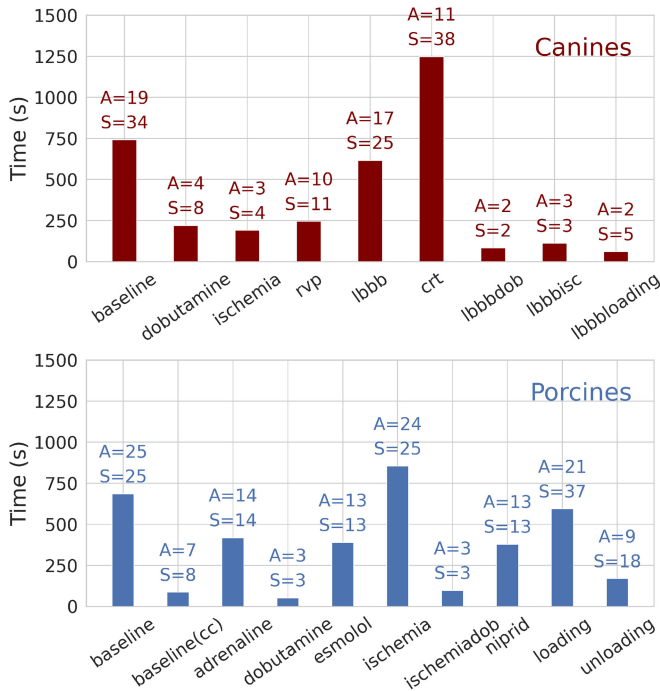
Fig. 2. Histogram showing the recorded time for each intervention of canines and porcines. The number of animals per intervention is given by A and the number of recordings is given by S. Some animals have multiple recordings per interventions to cover a larger data variety, e.g. an animal could have different degrees of loading.

ID: 8628 and 9303]. They were acute experiments where the animals were ventilated and surgically prepared as previously described [3], [27], [28]. The experiments followed different instrumentation protocols, but all of them had left ventricular pressure (LVP) and acceleration signals recorded simultaneously during different interventions. LVP was recorded using a calibrated micromanometer-tipped catheter (MPC-500, Millar Instruments Inc, Houston, TX). A tri-axial accelerometer sensor (MPU9250, InvenSense Inc, San Jose, CA, USA) was sutured to the epicardium in the LV apical, anterior region. The accelerometer's x-, y-, and z-axis were aligned with the longitudinal, circumferential, and radial directions, respectively. Depending on the protocol, data were recorded at either 650 Hz or 1000 Hz in canines and either 250 Hz or 500 Hz in porcines. The accelerometer sensor was calibrated to a unit of g.

B. Experimental Protocol

Fig. 2 shows the breakdown of interventions in the experiments. In the canine experiments, data were obtained from: baseline, right ventricular pacing (rvp), infusion of dobutamine, induction of ischemia, induction of left bundle branch block (LBBB) and subsequent bi-ventricular pacing for cardiac synchronization therapy (CRT). Data were also collected in a few animals combining LBBB with: infusion of dobutamine (lbbbdob), induction of ischemia (lbbbisc), or fluid loading (lbbbloading). Ischemia was induced by temporary occlusion of the proximal left anterior descending coronary artery (LAD). LBBB was induced by radio-frequency ablation of the left bundle

branch. Not all interventions were performed in all animals due to differences in protocols.

In the porcine experiments, data were obtained from two different experimental protocols. In the first case, data were taken from six different settings: baseline, infusion of adrenaline (epinephrine, 10 μg), infusion of beta-blocker (esmolol, 100 mg), infusion of vasodilator (niprid, 0.1 mg), ischemia induced as described above, and fluid loading. In the second set of porcine experiments, data were obtained during baseline, fluid loading, and phlebotomy (i.e. unloading). In some of the animals in the second set of porcine experiments, data were recorded both with open (represented as baseline) and closed chest (represented as baseline(cc)) during baseline. Furthermore, in three animals in the second set of experiments, data were recorded during infusion of dobutamine and infusion of dobutamine during ischemia (ischemiadob).

C. Annotation and Preprocessing

The LVP trace was passed through a smoothing window of 50 ms to remove any potential artifacts in the signal. The rate of change of the LVP (LV dP/dt) was then derived from the LVP signal. The time-differentiated signal was then passed through a smoothing window of 50 ms to remove any residual noise in the signal. Data labels were then automatically generated with the time point of maximum LV dP/dt taken as aortic valve opening (AVO) and the time point of minimum LV dP/dt taken as aortic valve closure (AVC) (Fig. 3). While minimum LV dP/dt is an established marker of AVC [29], we have not found a similar validation study of maximum LV dP/dt as a marker of AVO. Therefore, we investigated this in animals where also micromanometer measurements of aortic pressure (AoP) were available and the AoP catheter was positioned immediately proximal to the aortic valve to avoid transmission delay of the pressure wave to more distal positions in the aorta. AVO defined as the first point of rise in AoP, was compared to the time of maximum LV dP/dt. Maximum LV dP/dt occurred on average 5±7 ms (±SD) before the upstroke of AoP (Fig. 4) in 2500 heartbeats from three interventions in 17 canines. The average difference was less than 1% of the duration of the average heartbeat and thus, time of maximum LV dP/dt was considered an adequate label for AVO. While AoP recordings with verified proximal catheter position were available in a few of the animals, LVP was available in all animals. Hence, maximum LV dP/dt was used as the label for AVO. Finally, the LV dP/dt signal and generated data labels were manually checked and verified, where recording parts with false or missing AVO or AVC points were discarded.

The acceleration signals were re-sampled to a standard sampling rate of 500 Hz. The static gravity component of the acceleration signal was removed with a moving average filter (Tukey window of length 3 s with a cosine fraction of 0.5). To further reduce the variability of the input data, we used the magnitude of the acceleration only ($a_{magnitude} = \sqrt{a_x^2 + a_y^2 + a_z^2}$). Using the magnitude makes the approach insensitive to the orientation of the sensor axes, which is an advantage when attaching the
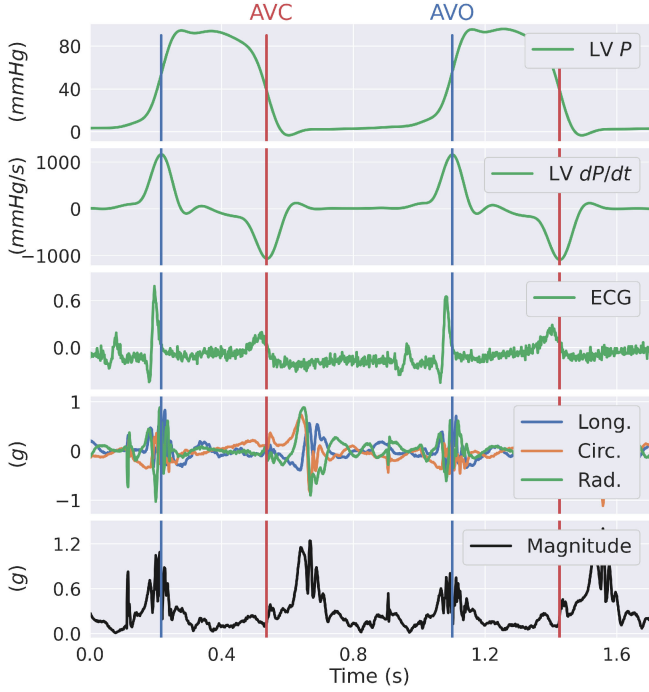
Fig. 3. Measured traces from a representative example where the valve events defined using LV *dP/dt* as the ground truth are shown. Onset Q in ECG coincide approximately with the onset of the systolic accelerometer waves in this case. This could be due to a fusion of a passive recoil of the ventricular wall following atrial induced filling with the active contraction starting the closing motion of the mitral valve leaflets. Long: longitudinal, Circ: circumferential, Rad: radial, AVO: aortic valve opening (blue), AVC: aortic valve closure (red).
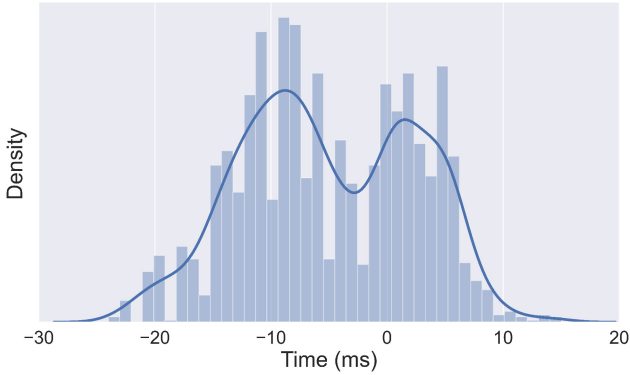

Fig. 4. Histogram showing the difference in timing of aortic valve opening by time-point of maximum LV *dP/dt* - first time-point of increase in aortic pressure.

sensor as it can be attached without concern for a specific orientation relative to the heart axes.

D. Proposed Deep Neural Networks

The cyclic motion of the heart with the sequential valve events, produces an acceleration signal with vibrations associated with different phases of the heart cycle, as can be seen in Fig. 3. The first module of the deep neural network was therefore a convolutional neural network (CNN) which is suited for pattern recognition of the typical vibrations associated with the valve

TABLE I
RESNET BLOCK

	Filter$_{in}$	Filter$_{out}$	Kernel=3	RC
Input ↓				
Conv 1D	C_{in}	C_{out}	3	
BatchNorm	-	-	-	
LeakyReLU	-	-	-	
Conv 1D	C_{out}	C_{out}	3	
BatchNorm	-	-	-	
LeakyReLU	-	-	-	
Conv	C_{in}	C_{out}	1	

The residual network (ResNet) block is used within the CNN module. The ResNet block can be seen as an activation before addition [30]. All layers use stride 1 and no dilation or padding. The residual connection is illustrated in the RC column.

TABLE II
CNN MODULE

	C_{in}	C_{out}	Kernel	Stride	Size(C x W)
Input ↓					1 x 1500
ResNet block	1	$32 \cdot k$	3	1	$(32 \cdot k)$ x 1496
Max pool	-	-	2	2	$(32 \cdot k)$ x 748
ResNet block	$32 \cdot k$	$64 \cdot k$	3	1	$(64 \cdot k)$ x 744
Max pool	-	-	2	2	$(64 \cdot k)$ x 372
ResNet block	$64 \cdot k$	$64 \cdot k$	3	1	$(64 \cdot k)$ x 368
Max pool	-	-	2	2	$(64 \cdot k)$ x 184
ResNet block	$64 \cdot k$	$128 \cdot k$	3	1	$(128 \cdot k)$ x 180

The table shows the CNN module for an input length of 1500 samples (3000 ms). The value of k used is 1. It is a hyper parameter which defines the number of filters in network. The CNN module is constructed to follow common best practices: Leaky ReLU, normalization layers, and skip connections.

TABLE III
RNN MODULE

	Hidden size	Layers	Bidirectional	Cell type
RNN	$2 \cdot 128 \cdot k$	1	Yes	GRU

The initial cell states were both trainable and initialized as a normal distribution with zero mean and unit variance.

events. The second module was a recurrent neural network (RNN) which is suited for connecting temporal features in relative close proximity in time. Finally, the third module was an attention module capable of connecting information across longer temporal distances. Fig. 5 shows the principle of the proposed method.

The CNN module can be viewed as sliding window filters along the time dimension, where the window size is given by the output neurons' receptive field. The receptive field (RF) was 68 samples (136 ms), and the stride (Δt) between the windows was 8 samples (16 ms). A schematic diagram of the neural network is shown in Fig. 6 and the corresponding modules are described in details in Tables I, II, III, IV, V. Let i denote the sliding window index. The network predicted: 1) whether an AVO $(\hat{t}_{AVO_i})$ and/or AVC $(\hat{t}_{AVC_i})$ were located within window i; 2) if the network predicted an event, the position of the AVO $(\hat{w}_{AVO_i})$ or the AVC $(\hat{w}_{AVC_i})$ within the window; and 3) the classification of species $(\hat{s}_i)$, porcine or canine.

(1) Gravity corrected acceleration magnitude

(2) Proposed deep neural network

(3) Combine network output

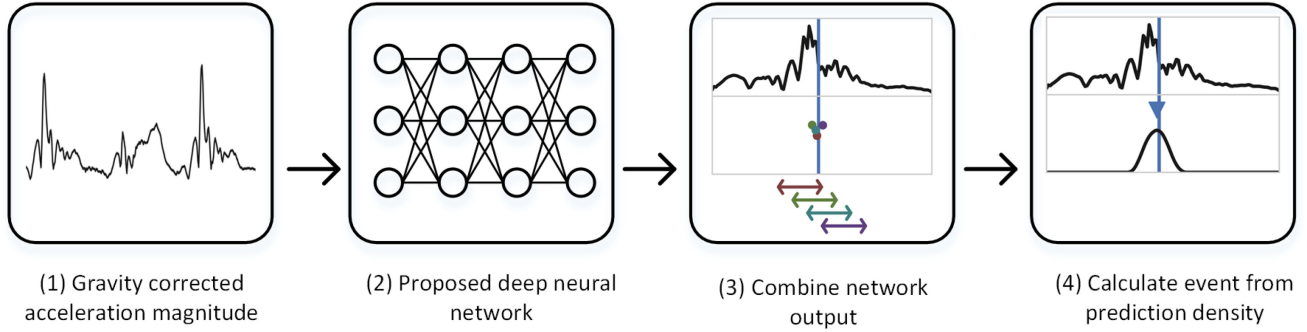(4) Calculate event from prediction density

Fig. 5. The figure shows a schematic diagram of the four stages in the proposed approach as explained in detail in the text. The output of the proposed neural network is predictions of the aortic valve opening and closure per window. Four such window estimates are plotted in (3) around an aortic valve opening event where the true target is shown as the blue vertical line. The windows and their corresponding estimates are color coded. The final event prediction (triangle) is calculated from the density of window estimates as shown in stage (4).
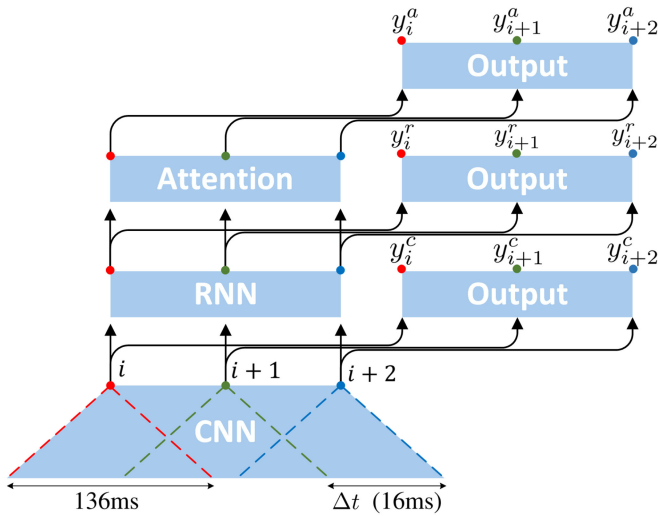

Fig. 6. A block diagram showing how the modules are connected in the neural network. The window size of 136 ms given by the CNN output neutrals receptive field is illustrated by the red, green, and blue dashed lines.

TABLE IV
ATTENTION MODULE

	C_{in}	Number of heads	p	RC
Input ↓				
PE	-	-	-	
MHA	$256 \cdot k$	8	-	
Dropout	-	-	0.5	
Input	-	-	-	

The multi-head self-attention (MHA) and positional encoding (PE) is defined as presented in [31]. The residual connection is illustrated in the RC column.

The prediction of species was added as an auxiliary task to regularize the network during training. Such regularization is important because when training is stopped at a given epoch, the network should perform well on the main predicted outputs $\hat{t}_{AVO_i}$, $\hat{t}_{AVC_i}$, $\hat{w}_{AVO_i}$, and $\hat{w}_{AVC_i}$. These four outputs were predicted for each module, and in addition the species prediction

TABLE V
OUTPUT MODULE

	C_{in}	C_{out}	kernel	p
Input ↓				
Conv 1D	$256 \cdot k$	$128 \cdot k$	1	-
LeakyReLU	-	-	-	-
Dropout	-	-	-	0.5
Conv 1D	$128 \cdot k$	5	1	-
Sigmoid	-	-	-	-

The output module connected to the CNN module have C_{in} of $128 \cdot k$, not $256 \cdot k$ as for the RNN and attention modules.

was a fifth output of the RNN and attention modules. The CNN module did not have enough context to predict the species on a per window level, and species was therefore not an output of this module. The estimates from the CNN and the RNN modules were included to improve the gradient flow. The output used for the final predictions was, however, only from the attention module.

With mathematical notation the prediction of the output targets can be described as follows. The superscripts c, r, and a are used to indicate the module references, CNN, RNN, and attention (ATT), respectively. Using u as a common term for the three superscripts, the output vector y_i^u of length five elements is:

$$\begin{bmatrix} y_i^u[0] = & \hat{t}_{AVO_i}^u \\ y_i^u[1] = & \hat{w}_{AVO_i}^u \\ y_i^u[2] = & \hat{t}_{AVC_i}^u \\ y_i^u[3] = & \hat{w}_{AVC_i}^u \\ y_i^u[4] = & \hat{s}_i^u \end{bmatrix}$$

Each output element is squeezed to a value between zero and one using the sigmoid function. The classification target t_{AVO_i} is one if window i includes an AVO event and zero otherwise. If an AVO event is located within window i, the regression target w_{AVO_i} denotes the normalized position of the AVO event within the window. The same structure applies to the AVC targets t_{AVC_i} and w_{AVC_i}. The target for the species predictions, (s_i),

is the same for each window i as it does not change during the recording.

The loss terms, $\mathcal{L}_{t_{AVO}}, \mathcal{L}_{t_{AVC}}, \mathcal{L}_s$, are computed using cross-entropy between the targets t_{AVO}, t_{AVC}, s and the network outputs $\hat{t}_{AVO}{}^u, \hat{t}_{AVC}{}^u, \hat{s}^u$. The regression losses are calculated using the L1 distance between the targets w_{AVO}, w_{AVC} and the network outputs $\hat{w}_{AVO}{}^u$ and $\hat{w}_{AVC}{}^u$ when events occur. The loss, $\mathcal{L}$, is a weighted linear combination of the five given losses $\mathcal{L}_{t_{AVO}}, \mathcal{L}_{w_{AVO}}, \mathcal{L}_{t_{AVC}}, \mathcal{L}_{w_{AVC}}$, and $\mathcal{L}_s$.

$$\mathcal{L}_{t_{AVO}} = \sum_{u=\{c,r,a\}} \sum_{i=1}^{N} H(t_{AVO_i}, \hat{t}_{AVO_i^u}) \quad (1)$$

$$\mathcal{L}_{w_{AVO}} = \sum_{u=\{c,r,a\}} \sum_{i=1}^{N} t_{AVO_i} \cdot |w_{AVO_i} - \hat{w}_{AVO_i^u}| \quad (2)$$

$$\mathcal{L}_{t_{AVC}} = \sum_{u=\{c,r,a\}} \sum_{i=1}^{N} H(t_{AVC_i}, \hat{t}_{AVC_i^u}) \quad (3)$$

$$\mathcal{L}_{w_{AVC}} = \sum_{u=\{c,r,a\}} \sum_{i=1}^{N} t_{AVC_i} \cdot |w_{AVC_i} - \hat{w}_{AVC_i^u}| \quad (4)$$

$$\mathcal{L}_s = \sum_{u=\{r,a\}} \sum_{i=1}^{N} H(s_i, \hat{s}_i^u) \quad (5)$$

$$\mathcal{L} = \lambda_t(\mathcal{L}_{t_{AVO}} + \mathcal{L}_{t_{AVC}})$$
$$+ \lambda_w(\mathcal{L}_{w_{AVO}} + \mathcal{L}_{w_{AVC}}) + \lambda_s \mathcal{L}_s \quad (6)$$

N is used to denote the total number of windows. With the example input size of 1500 samples (3000 ms) as given in Table II, N equals 180. The cross-entropy function is given as H. The weighing coefficients λ_t, λ_w, and λ_s are chosen such that the individual losses get the same order of magnitude. The values for λ_t, λ_w, and λ_s are 0.1, 1, 0.05, respectively.

E. Training and Augmentation

The total number of animals (46) is too small to perform the typical training, validation, and test data set split. A small test set would result in performance estimate with high variance. Therefore, we conducted a nested cross-validation with one inner fold and six outer folds. The data was split such that no recordings from an animal that was used for testing purposes had been seen by the network during training/validation. Furthermore, we trained and tested the network on canines' and porcines' data sets separately as well to see how that may alter the results. Fig. 7 illustrates how the folds were distributed. The validation data sets were used to perform early stopping based on the multitask loss, $\mathcal{L}$.

The length of each recorded sequence was different for the various interventions and animals. The typical range was between 15 s and 30 s, with a minimum of 5 s and a maximum of 100 s. During training, we randomly sampled short sub-sequences of length 3 s. This was done to reduce the likelihood of overfitting towards the longer sequences as the intra variability between heartbeats is smaller than inter variability between recordings.
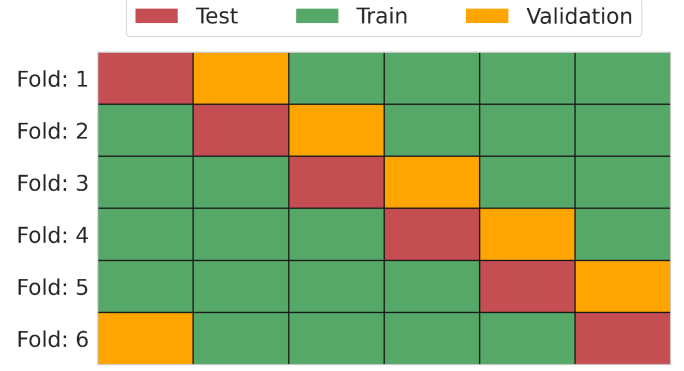

Fig. 7. Illustration of the nested cross-validation with one inner fold and six outer folds.

The sequences used for validation were of length 3 s, and the sequences used for testing were of maximum length 15 s.

The data was randomly augmented during training using time warping and magnitude scaling. The signals were stretched and gain adjusted with factors between 0.8 and 1.2.

F. Hyper-Parameter Search

A coarse hyper-parameter search was performed. The effect of different normalization layers such as batch normalization and group normalization were assessed as well as the effect of changing the model sizes through the value of $k = \{1, 2, 3, 4\}$ (Table II). As for the general network architecture, we report results from the individual output stages CNN, RNN, and attention (ATT) to determine their importance. Similarly, we evaluated the effect of not using a data driven localization estimate by setting w_{AVO} and w_{AVO} to the static value of 0.5. The max pool operation (with stride=2 and kernel=2) is not equivariant to translation. For this reason, we also experimented with Max blur pool (MBP) [32]. The use of the sigmoid function in the localization estimates has also been compared to the linear activation function (no sigmoid).

AdamW was used as optimizer [33]. The learning rates tested were 0.001 and 0.0001. The weight decay was fixed to 0.0001, and batch size was 32. The weights in the convolutional layers were initialized following Kaiming *et al.* [34], the hidden states in the RNN were initialed by sampling from a uniform distribution between $\pm 1/\sqrt{fan_{in}}$, and the weights in the multi-head attention were initialized following Xavier *et al.* [35].

G. Combining Network Output Into AVO and AVC Estimates

The output from the proposed deep learning network was AVO and AVC candidates for each window. As each window had length (RF) 136 ms and the stride between windows was 16 ms, an event had on average 8.5 overlapping windows. This produced up to a maximum of 9 candidate predictions for each event, and these predictions did not necessarily fall on the exact same sample but had a temporal distribution. The output of each window was therefore combined to give a final prediction as
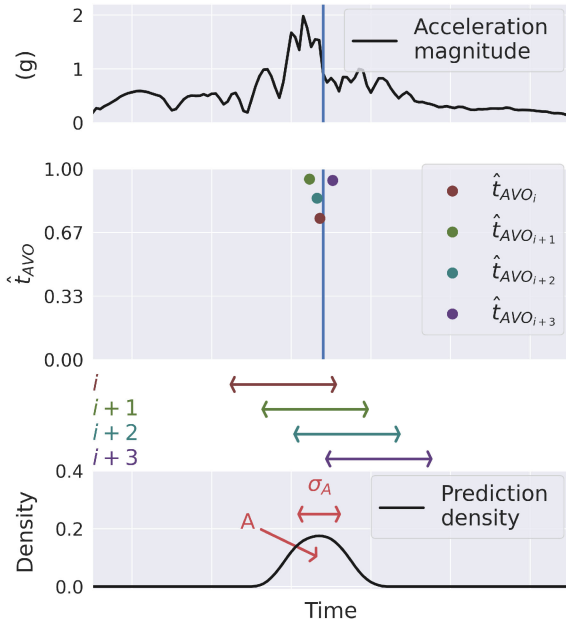
Fig. 8. First row: The acceleration magnitude is plotted together with an aortic valve opening (AVO) event. The second row illustrates four windows from running index i to $i + 3$. The windows are color coded to match their corresponding candidates. The third row shows the prediction density, with its area under the curve A, and its standard deviation σ_A.

described below for AVO only, as the steps were similar for AVC.

The estimated likelihood that window i contained an AVO was given by $\hat{t}_{AVO_i}$. The global location was calculated by $i \cdot \Delta t + RF \cdot \hat{w}_{AVO_i}$. Fig. 8 (second row) shows candidates from four windows around an AVO event. The estimated candidate likelihoods were re-scaled by,

$$\hat{s}_{AVO_i} = (\hat{t}_{AVO_i} - 0.5) \cdot 2, \tag{7}$$

and all re-scaled values ($\hat{s}_{AVO_i}$) smaller than zero were omitted. The re-scaling was performed to make the contribution from each window more representative. A candidate likelihood (e.g. $\hat{t}_{AVO_i}$) of 0.5 was considered to have zero confidence. A moving average filter of length 60 ms was then applied to generate a continuous prediction density curve, emphasizing regions with high number of AVO candidates and suppressing regions with few as illustrated in Fig. 8 third row.

A confidence score (C) with a corresponding position was calculated for each nonzero region in the prediction density. The confidence score of an AVO being present was given by $C = A/\sqrt{\sigma_A}$, and the position was calculated by the center of mass of A. Fig. 8 third row illustrates the predication density of a nonzero region with the metrics σ_A and A. The approach also included a three stage refinement process before reporting the final AVO estimates:

1) Confidence scores below 0.4 were rejected.
2) If two or more confidence scores were closer than 300 ms, the highest confidence score was kept only.

3) If an AVO confidence score was missing in between two consecutive AVC confidence scores, the highest AVO confidence score was reintroduced even if below 0.4.

The positions of the remaining confidence scores were the final AVO estimates.

H. Evaluation

To evaluate the performance of the method, predicted valve events were classified as correct or incorrect, and for the correct predictions, the time distance to its true label was assessed. To quantify whether AVO or AVC events had been detected, or falsely introduced, we defined a *detection distance limit* of 40 ms. If a final estimate lay within the duration of the *detection distance limit* from a label, the estimate was deemed as a correct detection. Correspondingly, if a final estimate was located with a distance larger than the *detection distance limit*, it was considered an incorrect detection. To assess the accuracy in milliseconds, the incorrect beats with detection outside the distance limit were excluded from this assessment as we wanted to evaluate how accurate the correct predictions were. Numbers from incorrect detections outside the *detection distance limit*, for example 100 or 200 ms away, would conceal the degree of accuracy of the correctly detected events and are considered of limited interest. The correct and incorrect detection rates, the mean absolute error (e_{mae}), and the root mean squared error (e_{rmse}) are reported per intervention and combined for all interventions.

As a smaller or larger *detection distance limit* would alter the correct and incorrect detection rates of the method, an analysis was performed to demonstrate the effect of varying the *detection distance limit* from 20 ms to 60 ms.

Due to the lack of temporal information from behind and in front of the current window, the approach is more likely to perform worse close to the borders of the time limited recordings. As the approach is to be used on continuous time series, AVO and AVC events located closer to the start and end than 300 ms were not included in the calculation of the results to avoid these potential higher errors.

We have used a gradient based optimization method for training the neural network. The seed to the pseudo random number generator (which influence among others the network's weight initialization) was shown to have a significant impact on the results. We quantified the sensitivity of the seed by performing nested cross-validation 50 times using different seeds. The reported results are from the hyper parameter configuration initialized with the particular seed that gave the lowest average error rate (AE) on the nested cross-validation data sets. The average error rate is defined as:

$$AE = \frac{1}{4} \sum_{\substack{s=\{canine, \\ porcine\}}} \sum_{\substack{e=\{AVO, \\ AVC\}}} ID_{e,s} + (100 - CD_{e,s}) \tag{8}$$

where ID is the incorrect detection rate and CD is the correct detection rate for all interventions.

We previously developed and tested an approach for detection of the valve events using a conventional signal processing
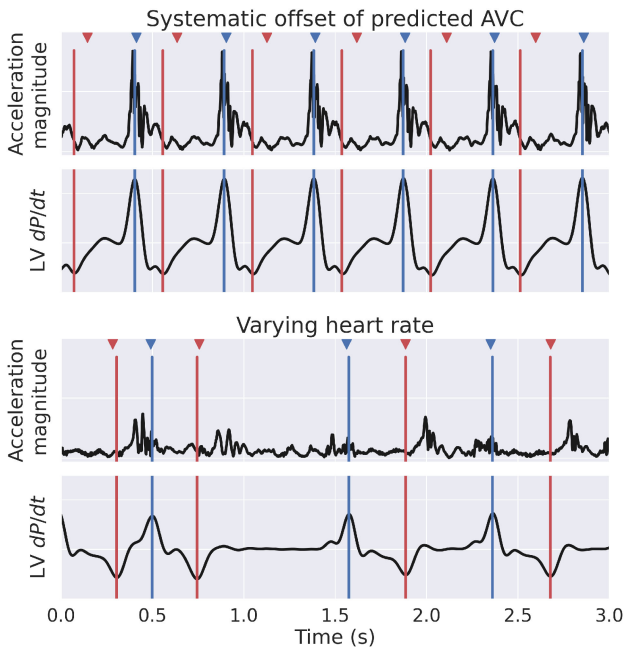
Fig. 9. The top panel from the adrenaline intervention in a porcine, illustrates a case with repeated failed detections of aortic valve closure (AVC) with a systematic error around 81 ms. The bottom panel from the ischemia intervention in a porcine, demonstrates a case with highly varying heart rate where aortic valve opening (AVO) and AVC are still detected with good accuracy for all beats. The blue and red triangles indicate AVO and AVC predictions, respectively, where the vertical lines are the corresponding references.

method [20]. The method employed different filter types (e.g. Butterworth, Chebyshev etc.) over specific frequency bands to emphasize the feature points associated with valve events on the acceleration signals. Basic peak/valley detection algorithms were then used to select these feature points associated with the valve events. The signal processing method was developed and validated on recordings from canines during only the three interventions: baseline, ischemia, and LBBB. We used this previous signal processing method to detect AVO and AVC on the data set of the current study including all interventions to compare the results with the neural network based approach.

III. RESULTS

The hyper parameters (as explained in II-F) that yielded the lowest average error rate given by (8) on the *validation data sets* was batch normalization, model size $k=1$, learning rate$=0.001$. The results reported in this section are on the *test data sets*.

The correct detection rates for AVO and AVC, pooling all interventions, were 98.9% and 97.1% in the canines and 98.2% and 96.7% in the porcines when defining a correct detection as within 40 ms of its true event. The mean absolute error between the correct detections and their corresponding targets was 8.4 ms and 7.2 ms for AVO and AVC in the canines, and 8.9 ms and 10.1 ms in the porcines.

The most common failure mode of the approach, was a systematic offset between the predictions and targets. This is not surprising as the signal is often repetitive. Fig. 9 top panels, illustrate this failure mode in case of the AVC. Irregular heart
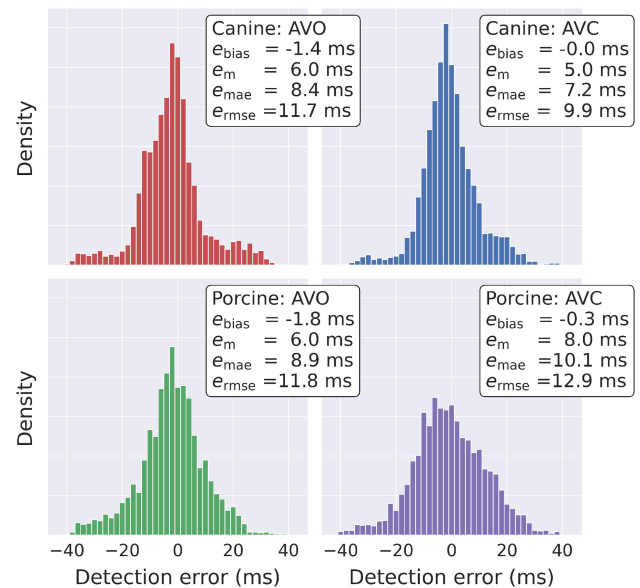

Fig. 10. Histograms showing the detection error in canines and porcines. The bias (e_{bias}), the median absolute error (e_m), the mean absolute error (e_{mae}), and the root mean square error (e_{rmse}) were calculated using a detection distance limit of 40 ms. AVO: aortic valve opening, AVC: aortic valve closure.

rhythm was not uncommon in several interventions with the duration of heartbeats varying with a factor of 2.6 in the extreme case as shown in bottom panels of Fig. 9. The method was therefore not restricted to rely on a consistent heart rate, and generally managed detection in cases with varying heart rates as shown in the figure.

The details of the performance for different interventions are given in Tables VI for canines and VII for porcines. The distribution of the detection errors is shown in Fig. 10. The figure shows that the majority of the tails of the distributions are within 40 ms. As can be inferred from the figure, the percentage of correct detections would decrease if a narrower *detection distance limit* was used to define the event as correct or incorrect. Correct and incorrect detection rates as a function of the *detection distance limit* in the range from 20 ms to 60 ms can be seen in Fig. 11. The figure also visualizes the variation in performance of using different seeds in the random number generator. The shaded areas in the figure show the standard deviation from the mean in the results from the 50 nested cross-validation runs. Table VIII displays the variation in performance in detail using a *detection distance limit* of 40 ms. The variation in true prediction was 2.6 percentage points in the worst case.

We tested the proposed method using individual species' data sets for training and testing. Using individual species' data sets yielded similar results to using merged data sets. Detailed results are available in the supplementary material.

Lastly, detection of the valve events by deep learning had higher feasibility than our previously proposed signal processing (SP) method. The pooled results for all interventions in both species are shown in Table IX. Several of the interventions caused alterations in the acceleration trace which increased the failing detection rate of the SP method.

TABLE VI
COMBINED RESULT FROM ALL TEST FOLDS - CANINES

Intervention	Number of animals	Number of recordings	AVO								AVC							
			Total number	Correct detection		Incorrect detection		Error [ms]		Total number	Correct detection		Incorrect detection		Error [ms]	
	#	#	#	#	%	#	%	e_{mae}	e_{rmse}	#	#	%	#	%	e_{mae}	e_{rmse}
baseline	19	34	710	708	99.7	2	0.3	8.9	12.9	717	710	99.0	7	1.0	7.9	11.6
dobutamine	4	8	172	172	100.0	0	0.0	4.2	5.1	171	171	100.0	0	0.0	5.2	6.0
ischemia	3	4	109	109	100.0	0	0.0	22.2	22.9	106	88	83.0	18	17.0	17.4	18.6
rvp	10	11	295	295	100.0	0	0.0	5.5	7.4	295	295	100.0	0	0.0	5.0	6.8
lbbb	17	25	538	537	99.8	1	0.2	11.9	14.6	544	513	94.3	31	5.7	8.2	11.2
crt	11	38	1101	1071	97.3	18	1.6	5.9	8.9	1098	1076	98.0	12	1.1	5.9	7.7
lbbbdob	2	2	64	63	98.4	0	0.0	10.5	10.9	66	65	98.5	0	0.0	8.9	9.7
lbbbisc	3	3	76	74	97.4	0	0.0	14.9	15.7	75	63	84.0	3	4.0	12.4	15.1
lbbbloading	2	5	98	98	100.0	0	0.0	8.6	9.8	99	98	99.0	1	1.0	7.4	8.8
all	19	130	3163	3127	98.9	21	0.7	8.4	11.7	3171	3079	97.1	72	2.3	7.2	9.9

Results for each intervention on canines. The calculations were done using a *detection distance limit* of 40 ms. The mean absolute error (e_{mae}) and the root mean square error (e_{rmse}) are calculated from the correct detections and the ground truth.

TABLE VII
COMBINED RESULT FROM ALL TEST FOLDS - PORCINES

Intervention	Number of animals	Number of recordings	AVO								AVC							
			Total number	Correct detection		Incorrect detection		Error [ms]		Total number	Correct detection		Incorrect detection		Error [ms]	
	#	#	#	#	%	#	%	e_{mae}	e_{rmse}	#	#	%	#	%	e_{mae}	e_{rmse}
baseline	25	25	550	550	100.0	0	0.0	8.4	11.2	558	555	99.5	1	0.2	9.5	12.1
baseline(cc)	7	8	139	139	100.0	0	0.0	4.3	5.8	133	128	96.2	4	3.0	10.7	13.1
adrenaline	14	14	439	428	97.5	0	0.0	10.6	12.9	430	393	91.4	24	5.6	10.7	13.1
dobutamine	3	3	108	107	99.1	0	0.0	10.6	12.6	107	106	99.1	0	0.0	8.0	9.2
esmolol	13	13	261	257	98.5	2	0.8	11.3	15.4	261	254	97.3	4	1.5	12.2	14.2
ischemia	24	25	569	564	99.1	3	0.5	8.4	10.9	567	561	98.9	4	0.7	9.8	12.6
ischemiadob	3	3	92	76	82.6	13	14.1	22.0	25.7	93	78	83.9	12	12.9	6.5	8.6
niprid	13	13	282	282	100.0	0	0.0	9.8	12.9	281	280	99.6	0	0.0	12.1	15.5
loading	21	37	620	602	97.1	16	2.6	6.7	8.4	602	589	97.8	11	1.8	9.0	11.8
unloading	9	18	293	288	98.3	3	1.0	7.6	10.4	281	261	92.9	17	6.0	11.4	14.6
all	27	159	3353	3293	98.2	37	1.1	8.9	11.8	3313	3205	96.7	77	2.3	10.1	12.9

Results for each intervention on porcine. The calculations were done using a *detection distance limit* of 40 ms. The mean absolute error (e_{mae}) and the root mean square error (e_{rmse}) are calculated from the correct detections and the ground truth.

IV. DISCUSSION

In this study, we have shown that the opening and closing of the aortic valve can be automatically detected by using deep neural networks on signals obtained through epicardially attached accelerometers. We trained and tested the network on data from a large set of interventions in canines and porcines. This was done to verify that the proposed approach is not restricted to a single species and works well under vastly varying cardiac motion and functional settings. The results support the concept that the continuous delineation of cardiac phases is possible from the accelerometer signal alone. This may improve the monitoring of cardiac function in cases where such an accelerometer is attached to the heart, for example accelerometers incorporated in CRT pace leads or the temporary pace leads that are routinely placed during open-heart surgery.

A main strength of our study, is the relatively large number of interventions from two different species, and the results which showed equally good performance from both animals in most cases using the deep learning approach. Several previously published methods to delineate valve events from SCG signals are feature- and ECG-dependent. Such features could include counting the number of peaks and troughs from the ECG R-peak or T-wave and the amplitude of peaks and troughs among others, which are then used in a decision tree based machine learning approach or signal processing method. This requires homogeneous data and may have limited accuracy, and thus limited clinical value under other settings where the signal morphology changes. We observed large variations in the signal as seen in Fig. 1, and our previous signal method [20] performed poorly in several of the tested interventions. On the other hand, the deep learning neural network detected the aortic valve events with high accuracy despite the large signal variation and with no use of ECG, demonstrating the ability of this method to identify patterns in such complex and varying signals. An additional strength of our study was our access to and use of
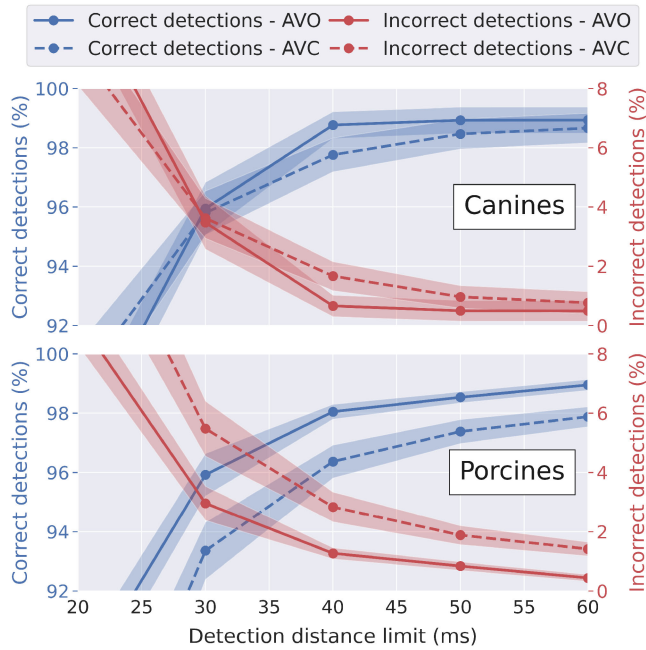
Fig. 11. Correct and incorrect detection rates as functions of the *detection distance limit*. The average (μ) of the 50 nested cross-validation runs using different seed to the random number generator, are given as lines, and the shaded regions correspond to the $\pm$ standard deviations (σ). AVO: aortic valve opening, AVC: aortic valve closure.

TABLE VIII
SENSITIVITY OF SEED

			μ	σ	Min	Max
Canines	AVO	CD	98.8	0.5	97.7	99.7
		ID	0.7	0.4	0.1	1.3
	AVC	CD	97.8	0.6	96.4	99.0
		ID	1.7	0.5	0.8	2.7
Porcines	AVO	CD	98.1	0.3	97.4	98.5
		ID	1.3	0.2	0.8	1.9
	AVC	CD	96.4	0.6	94.9	97.4
		ID	1.9	0.5	1.9	4.5

Summary of 50 nested cross-validation runs using different seeds in the random number generator and a *detection distance limit* of 40 ms. μ = mean, σ = standard deviation, CD = Correct detections, ID = Incorrect detections.

invasive LV pressure for reference of valve event time points. The pressure was measured simultaneously and synchronously with the accelerometer measurements, thus reducing subjectivity and errors in valve event annotations.

A. Network Structure

The suggested network follows a hierarchical structure. The backbone, the CNN module, is essential as it defines the local feature extractor, the window size, and the step length Δt. The window size of 136 ms was chosen such that the CNN module had adequate contextual information to perform independent predictions. The class imbalance of $\hat{t}_{AVO}$ and $\hat{t}_{AVC}$ is given by the ratio between the window size (136 ms) and the average heart cycle lengths. The average heart cycle length was 500 ms for

TABLE IX
COMPARISON BETWEEN SIGNAL PROCESSING (SP) AND NEURAL NETWORK (NN) BASED DETECTION OF AORTIC VALVE EVENTS

	AVO		AVC	
	Correct detection [%]	Error (e_{mae}) [ms]	Correct detection [%]	Error (e_{mae}) [ms]
SP	76.8	20.9	73.4	28.3
NN	98.5	8.6	96.9	8.7

Pooled results from all interventions in canines and porcines.

canines and 600 ms for porcines resulting in class imbalances. We typically seek to balance the classes which could be achieved by increasing the window size. However, we have observed in previous research that the detection error can increase with increased window size [22].

B. Hyper-Parameter Sensitivity Analysis

We experienced a large dependence on the selected seed to the random number generator. To trust the evaluation of different network configurations, we performed nested cross-validation with 50 different seeds (training 300 models per configuration). The result of our hyper-parameter sensitivity analysis on the *test data sets* is available in the supplementary materials. The attention module was included to yield global context to each window estimate. However, the results show no improvement over the RNN module. The performance increase with increased value of k. We did not find improved performance of using a data driven localization estimate compared to a static value of 0.5 in the case of the output stages RNN and ATT. However, the improvement is significant for the CNN output stage, where the window detector is less ideal. The confidence score threshold was reduced from 0.4 to 0.15 in the case of the static value of 0.5 to achieve a fair comparison to the data driven localization estimate.

We consider the max pool operation (with stride 2 and kernel size 2) to be problematic as it lacks translational equivariance. In an attempt to improve the data driven localization estimate, we tested Max blur pool (MBP). However, Max blur pool showed inferior performance to max pool.

Averaging over both species, canines and porcines, we did not find a performance difference between batch normalization and group normalization. However, interestingly, we found group normalization to favor canines, and batch normalization to favor porcines. For this reason, when applying this method on data from humans, the selection of normalization layer should the evaluated.

C. Results on Interventions

The neural network predicted the opening and closing of the aortic valve with reasonable accuracy on most of the interventions in both species, except for ischemiadob in porcines, and ischemia in canines. The lower accuracy in these interventions may have been attributed to the low number of recordings available for training and testing for these cases, or one or a few potentially bad recordings, or noisy data could result in lower

accuracy for the detected events as well. The acceleration traces and the distinct patterns associated with valve events changed entirely in cases where two interventions were combined i.e. infusion of dobutamine during ischemia (ischemiadob). This change in characteristic vibration pattern in the signal combined with the lower number of training data for these interventions could have contributed to a worse detection accuracy than other interventions.

D. Limitations

Our method is not without limitations. As opposed to SCG, which can be acquired non-invasively from the outside of the chest, our sensors are placed invasively on the heart for direct measurements of cardiac motion. Our method will be limited to patients where such high fidelity monitoring of cardiac function is required such as in cardiac surgery patients or patients in need of CRT. Another limitation of the proposed approach is that the data set is relatively small. Rather than having a traditional train, validation, and test split, we used nested cross validation to validate the accuracy of the proposed network. However, the small size of the data set is offset by the fact that the data were non-homogeneous with varying cardiac function and motion. Furthermore, the proposed method was developed in two different species which may pose a limitation as the cardiac motion and corresponding motion signal morphology may differ between the species. Our method may have to be modified and trained on human data before it can be applied in patients. However, we have previously directly translated a method for detection of ischemia from epicardial accelerometers in porcines to patients and obtained similar accuracy, indicating that myocardial motion is relatively similar [7]. Methods and data from chest-worn sensors in animals may be more difficult to translate to human use due to the additional difference in anatomy between the heart and skin including position of the heart within the chest. If the underlying cardiac motion is identical, such anatomical differences may alter the signal transmission to the skin and result in different signal morphology. Zia *et al.* [36] and Lin *et al.* [37] analyzed such non-invasive data from porcines and mentioned such a potential confounder. All methods must therefore be validated in humans before applications in patients.

The output from a neural network is only as good as the quality of its training data and correctly marked labels. Unfortunately, in some experiments the recorded LVP signal had artifacts that might have been introduced due to the pressure catheter touching the LV wall. These artifacts were further amplified when LV *dP/dt* was derived from the noisy LVP signals. Furthermore, the LV *dP/dt* signal frequently did not have a clear 'V-shaped' peak and trough for maximum LV *dP/dt* and minimum LV *dP/dt*, respectively, but instead had a 'W-shape' twin peak that would have made the marked AVO/AVC labels inconsistently even between consecutive beats. Therefore, the smoothing window of 50 ms was applied to both the LVP and LV *dP/dt* signals to remove these artifacts and to make the process of automatic label generation more robust and easier. This may have introduced inaccuracies in the labeling of the true events, but was still considered an improvement over using unfiltered LVP and LV

dP/dt signals to generate the labels. If the recorded LVP signals had been free of artifacts so no filtering/smoothing had to be applied to the signals, the mean error between the marked AVO and AVC labels and the predicted outputs would potentially be even lower.

Lastly, the results reported excludes AVO and AVC events closer to the beginning and the end of the sequence than 300 ms.

E. Future Work

This was a proof of concept study on canines and porcines. However, clinical studies need to be carried out to validate the proposed neural network on patient data. The same concept can be translated to accelerometer signals from humans. Furthermore, the deep learning approach to detect valve events can be further expanded to detect the opening and closing of the mitral valve as well. This may further improve monitoring of cardiac function by allowing more functional indices to be automatically calculated in real-time, such as the cardiac performance index (Tei index) which is the sum of the duration of the isovolumic contraction and relaxation periods divided by the duration of ejection. Lastly, a comparison of acceleration waveforms taken directly from the heart and SCG signals measured from the chest should be conducted, as it would lead to a better insight on the origin of SCG signals.

V. Conclusion

Deep neural networks can be used to automatically detect aortic valve opening and closing times using accelerometers attached to the heart. The proposed approach can handle a broad range of heart rates and does not require additional sensor inputs such as ECG. The method provided accurate and robust predictions on data from both porcines and canines, covering multiple interventions with varying cardiac motion and heart function. The results encourage translation of the method to the clinic for further investigations on how it can improve monitoring of cardiac function in patients.

Supplementary Material

Author Contributions

A.W., M.V.M, F.H.K., P.S.H, O.J.E. and E.W.R. participated in data collection. A.W. and T.S.J. conducted the data analysis and wrote the manuscript text together with E.W.R. All authors reviewed the manuscript.

Competing Interests

A.W., T.S.J., M.V.M., F.H.K., H.H.O., and A.H.S.S. report no conflict of interest. O.J.E. and P.S.H. are patent holders of the accelerometer technology for assessment of cardiac function,

and together with E.W.R. are share-holders in Cardiaccs A/S which is commercially exploiting cardiac accelerometers.

REFERENCES

[1] L. Chinitz et al., "Accelerometer-based atrioventricular synchronous pacing with a ventricular leadless pacemaker: Results from the Micra atrioventricular feasibility studies," *Heart Rhythm*, vol. 15, no. 9, pp. 1363–1371, 2018.

[2] T. Senoner et al., "Long-term performance of an atrial lead capable of accelerometer based detection of cardiac contractility in patients receiving cardiac resynchronisation therapy," *Plos One*, vol. 14, no. 9, 2019, Art. no. e0222269.

[3] P. S. Halvorsen et al., "Detection of myocardial ischaemia by epicardial accelerometers in the pig," *Brit. J. Anaesth.*, vol. 102, no. 1, pp. 29–37, 2009.

[4] A. Taebi, B. E. Solar, A. J. Bomar, R. H. Sandler, and H. A. Mansy, "Recent advances in seismocardiography," *Vibration*, vol. 2, no. 1, pp. 64–86, 2019.

[5] K. Sørensen, S. E. Schmidt, A. S. Jensen, P. Søgaard, and J. J. Struijk, "Definition of fiducial points in the normal seismocardiogram," *Sci. Rep.*, vol. 8, no. 1, pp. 1–11, 2018.

[6] M. E. Comunale et al., "The concordance of intraoperative left ventricular wall-motion abnormalities and electrocardiographic st segment changes: Association with outcome after coronary revascularization," *J. Amer. Soc. Anesthesiologists*, vol. 88, no. 4, pp. 945–954, 1998.

[7] P. S. Halvorsen et al., "Automatic real-time detection of myocardial ischemia by epicardial accelerometer," *J. thoracic Cardiovasc. Surg.*, vol. 139, no. 4, pp. 1026–1032, 2010.

[8] P. Brainin et al., "Presence of post-systolic shortening is an independent predictor of heart failure in patients following st-segment elevation myocardial infarction," *Int. J. Cardiovasc. Imag.*, vol. 34, no. 5, pp. 751–760, 2018.

[9] P. Brainin et al., "Post-systolic shortening predicts heart failure following acute coronary syndrome," *Int. J. Cardiol.*, vol. 276, pp. 191–197, 2019.

[10] T. Edvardsen, S. Urheim, H. Skulstad, K. Steine, H. Ihlen, and O. A. Smiseth, "Quantification of left ventricular systolic function by tissue doppler echocardiography: Added value of measuring pre-and postejection velocities in ischemic myocardium," *Circulation*, vol. 105, no. 17, pp. 2071–2077, 2002.

[11] R. C. Talley, J. F. Meyer, and J. L. McNay, "Evaluation of the pre-ejection period as an estimate of myocardial contractility in dogs," *Amer. J. Cardiol.*, vol. 27, no. 4, pp. 384–391, 1971.

[12] D. B. Newlin and R. W. Levenson, "Pre-ejection period: Measuring beta-adrenergic influences upon the heart," *Psychophysiol.*, vol. 16, no. 6, pp. 546–552, 1979.

[13] F. Khosrow-Khavar, K. Tavakolian, A. P. Blaber, J. M. Zanetti, R. Fazel-Rezai, and C. Menon, "Automatic annotation of seismocardiogram with high-frequency precordial accelerations," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 4, pp. 1428–1434, Jul. 2015.

[14] G. Shafiq, S. Tatinati, and K. C. Veluvolu, "Automatic annotation of peaks in seismocardiogram for systolic time intervals," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2016, pp. 2672–2675.

[15] T. Choudhary, L. Sharma, and M. K. Bhuyan, "Automatic detection of aortic valve opening using seismocardiography in healthy individuals," *IEEE J. Biomed. health Informat.*, vol. 23, no. 3, pp. 1032–1040, May 2019.

[16] I. Rivero, E. Valdes, and F. Valdes, "Robust detection of AO and IM points in the seismocardiogram using CWT," *IEEE Latin Amer. Trans.*, vol. 14, no. 11, pp. 4468–4473, Nov. 2016.

[17] F. Khosrow-Khavar, K. Tavakolian, A. Blaber, and C. Menon, "Automatic and robust delineation of the fiducial points of the seismocardiogram signal for noninvasive estimation of cardiac time intervals," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 8, pp. 1701–1710, Aug. 2017.

[18] H. K. Thakkar and P. K. Sahoo, "Towards automatic and fast annotation of seismocardiogram signals using machine learning," *IEEE Sensors J.*, vol. 20, no. 5, pp. 2578–2589, Mar. 2020.

[19] T. Choudhary, M. Bhuyan, and L. Sharma, "A novel method for aortic valve opening phase detection using SCG signal," *IEEE Sensors J.*, vol. 20, no. 2, pp. 899–908, Jan. 2020.

[20] A. Wajdan, M. R. Krogh, M. Villegas-Martinez, P. S. Halvorsen, O. J. Elle, and E. W. Remme, "Automatic detection of valve events by epicardial accelerometer allows estimation of the left ventricular pressure trace and pressure–displacement loop area," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, 2020.

[21] A. M. Fiorito, A. Østvik, E. Smistad, S. Leclerc, O. Bernard, and L. Lovstakken, "Detection of cardiac events in echocardiography using 3D convolutional recurrent neural networks," in *Proc. IEEE Int. Ultrasonics Symp.*, 2018, pp. 1–4.

[22] T. S. Jahren, E. N. Steen, S. A. Aase, and A. H. S. Solberg, "Estimation of end-diastole in cardiac spectral doppler using deep learning," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 67, no. 12, pp. 2605–2614, Dec. 2020.

[23] Y. Xiang, Z. Lin, and J. Meng, "Automatic QRS complex detection using two-level convolutional neural network," *Biomed. Eng. Online*, vol. 17, no. 1, pp. 1–17, 2018.

[24] M. Hammad, S. Zhang, and K. Wang, "A novel two-dimensional ECG feature extraction and classification algorithm based on convolution neural network for human authentication," *Future Gener. Comput. Syst.*, vol. 101, pp. 180–196, 2019.

[25] J. Yao, S. Tridandapani, W. Auffermann, C. Wick, and P. Bhatti, "An adaptive seismocardiography (SCG)-ECG multimodal framework for cardiac gating using artificial neural networks," *IEEE J. Transl. Eng. Health Med.*, vol. 6, pp. 1–11, 2018.

[26] C. Yang, B. D. Ojha, N. D. Aranoff, P. Green, and N. Tavassolian, "Classification of aortic stenosis using conventional machine learning and deep learning methods based on multi-dimensional cardio-mechanical signals," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, 2020.

[27] M. R. Krogh et al., "Continuous estimation of acute changes in preload using epicardially attached accelerometers," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 7, pp. 2067–2075, Jul. 2021.

[28] J. M. Aalen et al., "Mechanism of abnormal septal motion in left bundle branch block: Role of left ventricular wall interactions and myocardial scar," *JACC: Cardiovasc. Imag.*, vol. 12, no. 12, pp. 2402–2413, 2019.

[29] F. L. Abel, "Maximal negative DP/DT as an indicator of end of systole," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 240, no. 4, pp. H676–H679, 1981.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. IEEE Eur. Conf. Comput. Vis.* 2016, J. Leibe Matas, N. Sebe, and M. Welling, Eds. Cham: Springer International Publishing, 2016, pp. 630–645.

[31] A. Vaswani et al., "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.

[32] R. Zhang, "Making convolutional networks shift-invariant again," in *Proc. 36th Int. Conf. Mach. Learn.*, K. Chaudhuri and R. Salakhutdinov, Eds., 2019, vol. 97, pp. 7324–7334. [Online]. Available: https://proceedings.mlr.press/v97/zhang19a.html

[33] I. Loshchilov and F. Hutter, "Fixing weight decay regularization in Adam," 2018. [Online]. Available: https://openreview.net/forum?id=rk6qdGgCZ

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1026–1034.

[35] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, W. Teh and M. Titterington, Eds., Chia Laguna Resort, Sardinia, Italy, 2010, vol. 9, pp. 249–256. [Online]. Available: https://proceedings.mlr.press/v9/glorot10a.html

[36] J. Zia, J. Kimball, C. Rolfes, J.-O. Hahn, and O. T. Inan, "Enabling the assessment of trauma-induced hemorrhage via smart wearable systems," *Sci. Adv.*, vol. 6, no. 30, 2020, Art. no. eabb 1708.

[37] D. J. Lin, J. P. Kimball, J. Zia, V. G. Ganti, and O. T. Inan, "Reducing the impact of external vibrations on fiducial point detection in seismocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 1, pp. 176–185, Jan. 2022.