Title

Identification of supraventricular tachycardia mechanisms with surface electrocardiograms using a convolutional neural network.

Journal

Heart Rhythm O2, 4(8)

Authors

Higuchi, Satoshi
Li, Roland
Gerstenfeld, Edward
et al.

Publication Date

2023-08-01

DOI

10.1016/j.hroo.2023.07.004

Peer reviewed

Identification of supraventricular tachycardia mechanisms with surface electrocardiograms using a convolutional neural network

Satoshi Higuchi, MD,[*1] Roland Li, BA,[†1] Edward P. Gerstenfeld, MD, FHRS,[*] L. Bing Liem, DO, FHRS,[*‡] Sung Il Im, MD,[*] Shadi Kalantarian, MD, MPH,[*] Minhaj Ansari, MSc,[†] Sean Abreau, MSc,[†] Joshua Barrios, PhD,[†] Melvin M. Scheinman, MD, FHRS,[*] Geoffrey H. Tison, MD, MPH[†§]

From the [*]Section of Cardiac Electrophysiology, Division of Cardiology, University of California, San Francisco, San Francisco, California, [†]Division of Cardiology, Department of Medicine, University of California, San Francisco, San Francisco, California, [‡]Division of Cardiology, San Francisco VA Medical Center, San Francisco, California, and [§]Bakar Computational Health Sciences Institute, University of California, San Francisco, San Francisco, California.

BACKGROUND It remains difficult to definitively distinguish supraventricular tachycardia (SVT) mechanisms using a 12-lead electrocardiogram (ECG) alone. Machine learning may identify visually imperceptible changes on 12-lead ECGs and may improve ability to determine SVT mechanisms.

OBJECTIVE We sought to develop a convolutional neural network (CNN) that identifies the SVT mechanism according to the gold standard of SVT ablation and to compare CNN performance against experienced electrophysiologists among patients with atrioventricular nodal re-entrant tachycardia (AVNRT), atrioventricular reciprocating tachycardia (AVRT), and atrial tachycardia (AT).

METHODS All patients with 12-lead surface ECG during sinus rhythm and SVT and had successful SVT ablation from 2013 to 2020 were included. A CNN was trained using data from 1505 surface ECGs that were split into 1287 training and 218 test ECG datasets. We compared the CNN performance against independent adjudication by 2 experienced cardiac electrophysiologists on the test dataset.

RESULTS Our dataset comprised 1505 ECGs (368 AVNRT, 304 AVRT, 95 AT, and 738 sinus rhythm) from 725 patients. The CNN areas under the receiver-operating characteristic curve for AVNRT, AVRT, and AT were 0.909, 0.867, and 0.817, respectively. When fixing the specificity of the CNN to the electrophysiologist adjudicators' specificity, the CNN identified all SVT classes with higher sensitivity: (1) AVNRT (91.7% vs 65.9%), (2) AVRT (78.4% vs 63.6%), and (3) AT (61.5% vs 50.0%).

CONCLUSION A CNN can be trained to differentiate SVT mechanisms from surface 12-lead ECGs with high overall performance, achieving similar performance to experienced electrophysiologists at fixed specificities.

KEYWORDS Convolutional neural network; Machine learning; Artificial intelligence; Supraventricular tachycardia; Electrocardiogram; Long RP tachycardia

(Heart Rhythm O^2 2023;4:491–499)

Introduction

Catheter ablation is the treatment of choice for most patients with supraventricular tachycardias (SVTs).[1–3] It is important for clinicians to be able to discern the underlying mechanism of SVTs including atrioventricular nodal re-entrant tachycardia (AVNRT), atrioventricular reciprocating tachycardia (AVRT), and atrial tachycardia (AT) in order to estimate chances of ablation success and possible adverse effects. The 12-lead electrocardiogram (ECG) is the most common diagnostic modality for preprocedural assessment. Current clinical paradigms emphasize the importance of both the morphology and relationship of the P-wave to the QRS during tachycardia in order to establish the SVT diagnosis.[4–12] However, existing 12-lead ECG paradigms can result in overlap in distinguishing between these SVT mechanisms.

Recent work has shown that machine learning can be effectively applied to ECGs to identify well-understood as well as novel ECG-based diagnoses.[13–19] Machine learning algorithms such as convolutional neural networks (CNNs)

[1]The first two authors contributed equally to this article. **Address reprint requests and correspondence:** Dr Geoffrey H. Tison, Division of Cardiology, Department of Medicine, Bakar Computational Health Sciences Institute, University of California, San Francisco, 555 Mission Bay Boulevard, South Box 3120, San Francisco, CA 94158. E-mail address: geoff.tison@ucsf.edu; Twitter: @GeoffTison

2666-5018/

https://doi.org/10.1016/j.hroo.2023.07.004

KEY FINDINGS

- In this study, we trained a convolutional neural network (CNN), a type of artificial intelligence algorithm, to analyze raw 12-lead electrocardiogram waveforms to identify 2 supraventricular tachycardia classes according to a gold standard of electrophysiology study diagnosis: atrioventricular nodal re-entrant tachycardia, atrioventricular reciprocating tachycardia, and atrial tachycardia.
- The CNN area under the receiver-operating characteristic curve for atrioventricular nodal re-entrant tachycardia, atrioventricular reciprocating tachycardia, and atrial tachycardia were 0.909, 0.867, and 0.817, respectively.
- The CNN performed similarly to cardiac electrophysiologists who adjudicated supraventricular tachycardia using 12-lead surface electrocardiograms alone, achieving similar sensitivities to electrophysiologists at fixed specificities.

allow for providing a valuable data-driven complement to human-defined rules that consider only a fraction of the total available ECG data. Therefore, the focus of the current study was to determine whether the use of machine learning analysis of surface 12-lead ECGs alone can accurately determine arrhythmia mechanisms of SVT as adjudicated by intracardiac recordings during an electrophysiology study. In this study, we developed and trained a CNN to discriminate between the SVTs of AVNRT, AVRT, and AT against the gold-standard SVT mechanism diagnosis derived from successful electrophysiology ablation studies for each patient. We then compared CNN performance to discriminate SVT mechanisms against independent adjudication by 2 experienced cardiac electrophysiologists.

Methods

Study population

We retrospectively reviewed 981 consecutive patients who underwent an electrophysiology study and catheter ablation at UCSF Medical Center between January 2013 and January 2020. Among them, a total of 256 patients were excluded: 148 patients failed to have inducible sustained SVT; 34 had noise due to bad electrode contacts, motion artifacts, and electromyography noise; 9 had missing leads; 12 had multiple tachycardia mechanisms or uncertain SVT diagnosis; and 21 had rare SVT forms including 3 nodo-ventricular/fascicular accessory pathways, 7 antidromic reciprocal tachycardias, 2 junctional tachycardias, and 9 permanent junctional reciprocating tachycardias. Moreover, 32 patients with ATs or AVNRTs with a 2:1 atrioventricular ratio were also excluded, as those forms can exclude AVRT with this finding alone. A total of 725 patients were included in the study analysis (Figure 1). This study was conducted according to the principles of the Declaration of Helsinki, and the study protocol was approved by the University of California San Francisco Institutional Review Board. All electrophysiology studies were initially performed with a signed informed consent; the need for informed consent for the research analysis was waived by the Institutional Review Board in the setting of anonymized retrospective record review.

Electrophysiology study

Twelve-lead surface ECGs and intracardiac ECGs were recorded and stored on the Prucka CardioLab (GE Healthcare, Waukesha, WI) recording system. Twelve-lead ECGs were filtered between 0.05 and 100 Hz and bipolar intracardiac electrograms between 30 and 500 Hz and recorded at a speed of 100 mm/s. Tachycardia cycle length, VA interval, and QRS interval during SVTs were measured in each patient using digital calipers of the intracardiac electrogram. The

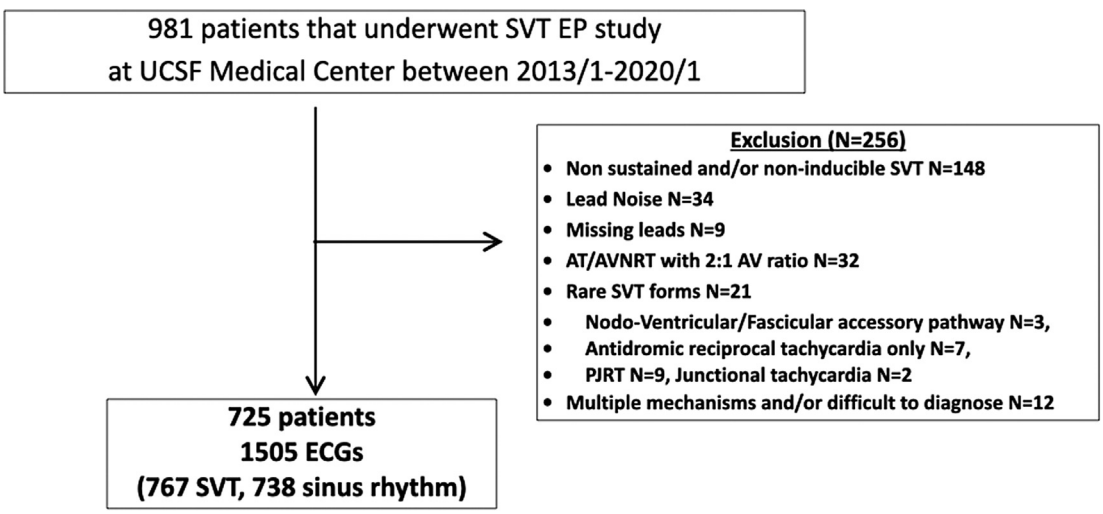

Figure 1 Diagram of study electrocardiogram (ECG) datasets. AT = atrial tachycardia; AVNRT = atrioventricular nodal re-entrant tachycardia; EP = electrophysiology; PJRT = permanent junctional reciprocating tachycardia; SVT = supraventricular tachycardia; UCSF = University of California, San Francisco.

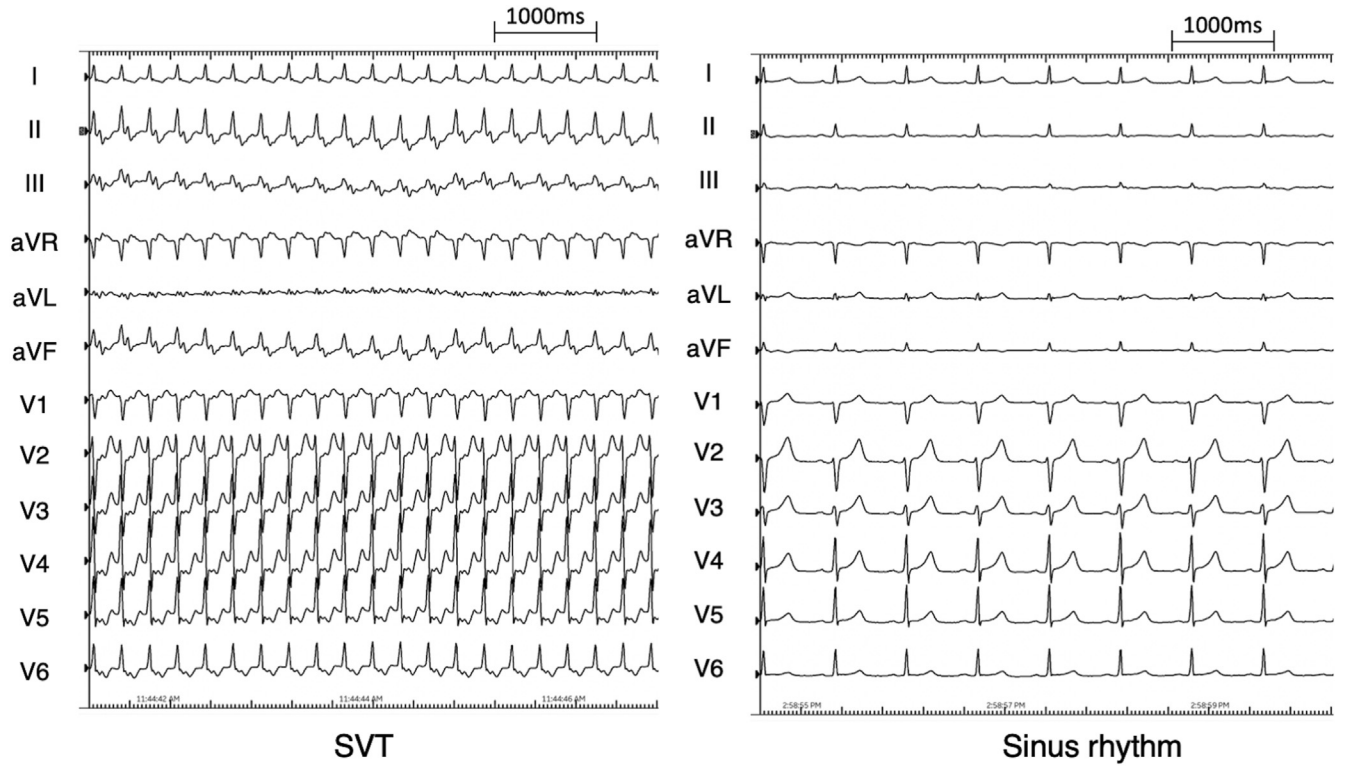

Figure 2 Adjudication by the expert electrophysiologists. Twelve-lead electrocardiograms of supraventricular tachycardia (SVT) and sinus rhythm (as a reference) were provided to ask the most likely SVT mechanism among atrioventricular nodal re-entrant tachycardia, atrioventricular reciprocating tachycardia, atrial tachycardia, or undetermined.

diagnoses of 3 SVT mechanisms were confirmed according to (1) conventional electrophysiological criteria using right ventricular overdrive pacing, right ventricular extrastimuli, and atrial extrasimuli during tachycardia; and (2) the results of a successful ablation site.[20–23] All electrophysiological studies and ablation studies were reviewed by electrophysiologists (S.H., S.I.I.), and the SVT mechanisms were reconfirmed again by a senior electrophysiologist (M.M.S.).

Datasets and ECG data

All raw ECG waveform data were extracted from the Prucka CardioLab recording system. Ten seconds of surface 12-lead ECGs during SVT were extracted from all patients. Moreover, 10 seconds of surface 12-lead ECGs during sinus rhythm were also extracted from the same patients in order to provide data for an ancillary task of sinus rhythm classification, which improves algorithm training for SVT identification. The extracted ECG data included only those during SVT or sinus rhythm; any other rhythms were excluded. All ECG data were split into training (1287 [sinus rhythm: 630, AVNRT: 308, AT: 82, AVRT: 267]) and test (218 [sinus rhythm: 108, AVNRT: 60, AT: 13, AVRT: 37]) datasets, with each dataset containing nonoverlapping patients. During CNN training, to address class imbalance, minority classes were upsampled randomly with replacement to a ratio of sinus:AVNRT:AT:AVRT of 1:0.75:0.75:0.75. The CNN was trained and evaluated using labels of SVT mechanisms defined by the electrophysiology study and successful catheter ablation.

Adjudication by cardiac electrophysiologists

An adjudication dataset was identified, which contained a total of 200 SVT examples drawn from the full dataset. All 110 SVT examples in the test dataset (including 60 AVNRT, 37 AVRT, and 13 AT) were included in the adjudication dataset, along with 90 other SVT examples randomly selected from the training dataset. Two expert electrophysiologist physicians (E.P.G., L.B.L.) used standard surface ECG approaches or criteria to independently adjudicate the SVT mechanism (AVNRT, AVRT, AT, or undetermined) from the surface 12-lead ECG alone (see Discussion).[4–12] This provided cardiac electrophysiologist SVT adjudication performance to compare against CNN performance for the entire test dataset, along with a larger adjudicated dataset for long RP SVT substrata, which are more rare but clinically relevant. In each dataset, sinus rhythm 12-lead ECGs from each patient were also extracted (Figure 2).

CNN algorithm architecture

ECG lengths were standardized to the 75th percentile of ECG lengths in our dataset (19,532 sample voltages) by either cropping or zero-padding. ECGs voltage values were normalized per lead per ECG via the Yeo-Johnson power transform to achieve a gaussian distribution around zero of the deep neural network input.[24] CNN models were trained

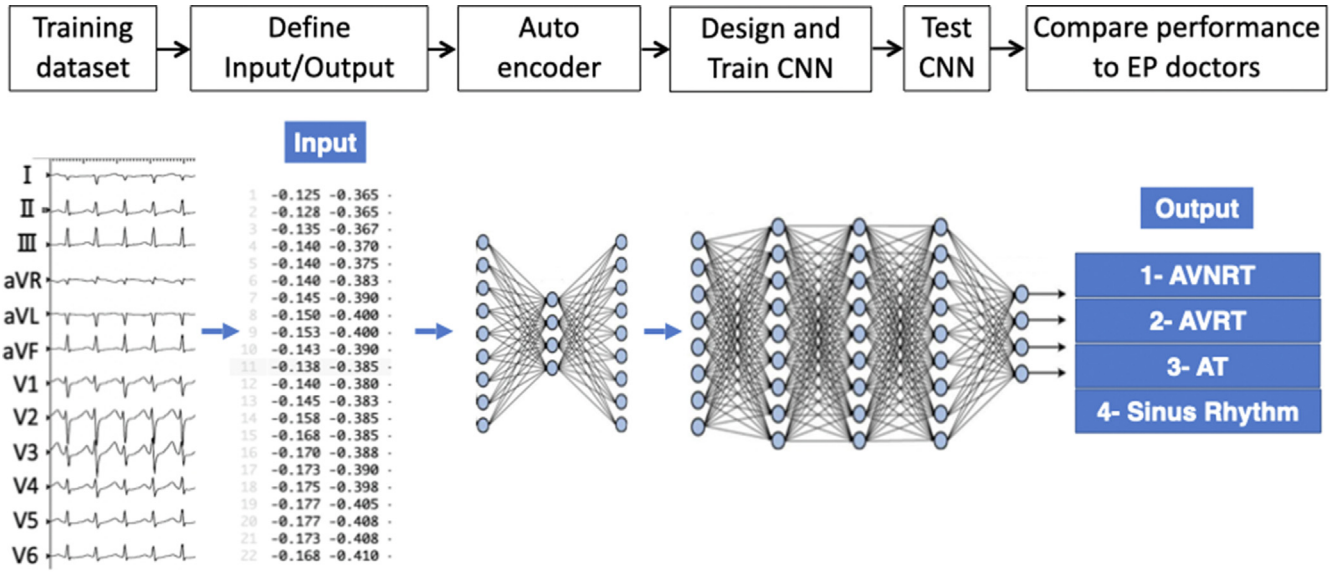

Figure 3 Convolutional neural network (CNN) model development. AT = atrial tachycardia; AVNRT = atrioventricular nodal re-entrant tachycardia; AVRT = atrioventricular reciprocating tachycardia; EP = electrophysiology.

to predict the presence of each of the 4 diagnostic classes: AVNRT, AVRT, AT, and sinus rhythm. For CNN training, sinus rhythm was included as a fourth non-SVT class as an ancillary task, which served to improve overall CNN performance on the 3 target SVT classes of primary interest. In the first model, an autoencoder,[25] took all ECGs, each sized 19,532 × 12, and extracted embeddings of size 611 × 128 for each ECG (Figure 3). An autoencoder reduced input data into smaller embeddings that retained important features and learned to recreate the original input from those embeddings. These embeddings were then passed into a 1-dimensional ResNet-based CNN architecture similar to that previously described.[15,16] The outputs of this CNN were 4 continuous predictions between 0 and 1 corresponding to probability for each diagnosis. The model consisted of 13 stacked convolutional layers with a filter size of 11, each with a ReLU nonlinearity layer[26] and batch normalization layer.[27] For every second convolutional layer, a max-pooling layer was applied, reducing the time resolution by a factor of 2 starting from 611 and every other a residual connection. The number of convolutional channels was doubled every fourth layer starting from 128. The final hidden dense layer had a 128 × 1 shape and was fed into a softmax activation layer to output probabilities of the 4 SVT classes (Figure 3). Each CNN was trained to converge the categorical cross-entropy loss function with an Adam optimizer and an initial learning rate of 10^{-3}, reduced by a factor of 0.5 every 5 epochs with no significant reduction in validation loss.[28] The optimal initial learning rate, dropout, and learning rate decay patience and factor and the size and

Table 1 Demographic and Diagnosis Characteristics

		AVNRT	AVRT	AT	*P* values
1. Patients (N=725)*		**356**	**292**	**81**	
Age, y		45 ± 22	22 ± 16	50 ± 22	<.001
Female		203 (57)	118 (40)	47 (58)	<.001
Long RP tachycardia		34 (10)	26 (9)	62 (77)	
Short RP tachycardia		322 (90)	266 (91)	19 (23)	
2. ECGs (N=767)		**368**	**304**	**95**	
Mechanisms		Typical form 89% Atypical form 11%	Left FW AP 52% Septal AP 36% Right FW AP 12%	RA AT 58% Septal AT 26% LA AT 16%	
12-lead ECG presentation	Bundle branch block (QRS width ≧120ms)	37 (10)	41 (13)	13 (14)	.36
Intracardiac ECG findings	Tachycardia cycle length, ms	354 ± 75	318 ± 58	397 ± 83	<.001
	VA interval during SVT, ms	43 ± 64	104 ± 37	194 ± 93	<.001
	QRS interval during SVT, ms	83 ± 26	82 ± 19	85 ± 21	.47

All data are expressed as the mean ± SD or n (%). AVNRT = atrioventricular nodal re-entrant tachycardia, AVRT = atrioventricular reciprocating tachycardia, AT = atrial tachycardia, SVT = supraventricular tachycardia, FW = free wall, AP = accessory pathway, RA =right atrial, LA = left atrial.
*Some patients contributed more than one disease record and ECG.

number of hidden layers were all found using a grid search across values consistent with those used in previous work. We used Keras version 2.2.4 [29] and Python 3.7 to train the models using an NVIDIA GTX 1080 Ti with 12 GB of VRAM (NVIDIA Corporation, Santa Clara, CA).[30]

Statistical analysis

We evaluated our CNN performance using the area under the receiver-operating characteristic curve (AUC) on the respective test datasets. We calculated bootstrapped confidence intervals for all metrics. We evaluated our CNN using binary decision thresholds for each diagnostic class chosen based on the optimal F1 score. For the test dataset, we reported other metrics including the positive predictive value (PPV), negative predictive value (NPV), sensitivity, specificity, and F1 score at these F1 score optimized thresholds. When comparing the CNN with the SVT ECGs adjudicated by cardiac electrophysiologists, for each class separately, we changed the CNN threshold such that the CNN specificity approximately matched the average specificity of the 2 electrophysiologist adjudicators for that class. This allows comparison of relative sensitivities between the CNN and expert electrophysiologist for each class separately.

Results

Study population

A total of 1505 (767 SVT and 738 sinus rhythm) ECGs were obtained (Figure 1). Table 1 summarizes patient age and sex based on the SVT mechanism and SVT characteristics evaluated from the intracardiac ECG. Patients with AVNRT and AT were older and more likely to be female, while patients with AVRT were younger and more likely to be male. The ECGs with AVNRT included 89% with a typical form and 11% with an atypical form. The ECGs with AVRT were most frequently left free wall accessory pathways (52%), followed by septal pathways (36%) and then right free wall pathways (12%). Right-sided AT was found in 58% of the AT ECGs, followed by septal AT (26%) and then left-sided AT (16%). Around 90% of AVNRT and AVRT ECGs presented with short RP tachycardia, while 77% of AT ECGs presented with long RP tachycardia forms. No significant difference in the prevalence of bundle branch block was found among the 3 SVT mechanisms ($P = .36$). For the detailed ECG characteristics, the tachycardia cycle length was shortest in those with AVRT (318 ± 58 ms) and the VA interval was shortest in those with AVNRT (43 ± 64 ms) followed by AVRT (104 ± 37 ms) and then AT (194 ± 93 ms).

Algorithm performance

Among the 3 SVT mechanisms, the CNN demonstrated the highest AUC in the test dataset for AVNRT (CNN AUC 0.909, 95% confidence interval [CI] 0.874–0.934), followed by AVRT (CNN AUC 0.867, 95% CI 0.796–0.913), and then AT (CNN AUC 0.817, 95% CI 0.739–0.891). CNN sensitivity and specificities in Table 2 are shown at thresholds optimized by F1 score, resulting in relatively higher specificities

Table 2 Performance of the convolutional neural network to differentiate the SVT mechanisms by 12-lead electrocardiogram in the test dataset

	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	F1 (95% CI)
AVNRT	0.909 (0.874–0.934)	80.9% (0.724–0.877)	87.9% (0.830–0.922)	72.5% (0.600–0.804)	92.7% (0.891–0.951)	76.0% (0.672–0.826)
AVRT	0.867 (0.796–0.913)	46.7% (0.354–0.625)	94.9% (0.917–0.973)	65.5% (0.498–0.801)	90.2% (0.859–0.931)	54.4% (0.418–0.667)
AT	0.817 (0.739–0.891)	45.8% (0.230–0.669)	95.1% (0.917–0.971)	36.0% (0.167–0.547)	96.5% (0.941–0.985)	41.0% (0.188–0.581)
Sinus rhythm*	0.986 (0.967–0.999)	97.1% (0.942–1.000)	98.0% (0.958–1.000)	97.9% (0.954–1.000)	97.1% (0.940–1.000)	97.4% (0.954–0.989)

AT = atrial tachycardia; AUC = area under the receiver-operating characteristic curve; AVNRT = atrioventricular nodal re-entrant tachycardia; AVRT = atrioventricular reciprocating tachycardia; CI = confidence interval; NPV = negative predictive value; PPV = positive predictive value; SVT = supraventricular tachycardia.

*For convolutional neural network training, sinus rhythm was included as a fourth non-SVT class as an ancillary task, which served to improve overall convolutional neural network performance on the 3 target SVT classes of primary interest. Convolutional neural network performance for sinus rhythm is presented here for completeness.

and NPVs for all SVT classes. The CNN threshold can be altered according to the target clinical application to favor either higher sensitivity or specificity. For sensitivity analysis, we performed a 10-fold cross-validation whereby the CNN was trained and evaluated on 10 different randomly split training/development/test datasets, and averaged results across the 10 folds were not materially changed.

Two experienced cardiac electrophysiologists adjudicated all 110 SVT examples in the test dataset (Table 3). The CNN threshold was then changed to fix CNN specificity at approximately the (relatively high) average specificity of the cardiac electrophysiologists. This enabled the comparison of sensitivity between CNN and electrophysiologists for each SVT class separately. At those fixed specificities, the sensitivity of the CNN was higher than the averaged electrophysiologist sensitivity for all SVT classes: AVNRT (91.7% vs 65.85%), AVRT (78.4% vs 63.6%), and AT (61.5% vs 50.0%) (Table 3, asterisks).

CNN performance was then calculated for strata of long and short RP SVT tachycardia in the test dataset (Table 4). Such stratification is informative because long RP tachycardias are typically more challenging to identify and have fewer 12-lead ECG characteristics by which to discern the respective mechanisms. Due to the small sample sizes in respective substrata, AVRT in long RP strata and AT in short RP strata are excluded from stratified analysis. CNN performance in long RP tachycardia strata was lower compared with Table 2, though it remained moderately strong. CNN performance was similar for the short RP tachycardia strata compared with the unstratified analysis.

Because long or short RP stratification decreased sample sizes in substrata, as a sensitivity analysis we compared the performance of the CNN vs cardiac electrophysiologists in all available electrophysiologist adjudicated data of 200 SVT examples (including all 110 in the test dataset). The CNN threshold was fixed at the average specificity of the cardiac electrophysiologists allowing comparison of sensitivity among strata of long and short RP tachycardia (Table 5). At those specificities, the sensitivities of the CNN were higher for both long RP tachycardia (AVNRT and AT) and short RP tachycardia (AVNRT and AVRT) compared with the electrophysiologists. It is notable that CNN performance demonstrated relatively high specificity in both AVNRT and AT long RP strata, which is clinically difficult to identify.

Discussion

This work demonstrates that CNN analysis of surface ECGs achieves strong overall performance (AUC >0.8) to discriminate all 3 SVT mechanisms against the gold standard of electrophysiology study and catheter ablation, achieving similar performance using surface 12-lead ECG to experienced cardiac electrophysiologists but in an automated manner. At fixed specificities approximately similar to experienced electrophysiologists, the CNN exhibited higher sensitivity for AVNRT, AVRT, and AT. At the CNN thresholds used, the CNN had higher NPV and lower PPV compared with average physician performance. These results demonstrate that SVT subclasses can be distinguished using readily available surface 12-lead ECGs analyzed by a CNN artificial intelligence algorithm, achieving comparable performance to expert electrophysiology physicians and offering a potential complement to existing clinical paradigms to discern underlying SVT mechanisms.

SVTs denote tachyarrhythmias that originate from supraventricular tissue (AVNRT or AT) or require it to be a part of a re-entrant AV circuit (AVRT). AVNRT tends to be the predominant SVT mechanism (56%), followed by AVRT (27%) and then AT (17%).[3] Noninvasive differentiation of the underlying SVT mechanism using surface 12-lead ECG is an important clinical task, with the potential to assist in counseling and determining the procedural complication risk and facilitate the electrophysiology study planning and ablation strategy, and may reduce the procedure duration as well as radiation time. Current clinical paradigms that aim to ascertain SVT mechanism by 12-lead ECGs emphasize the importance of both the morphology and relationship of the

Table 3 Performance of the experienced cardiac electrophysiologists vs CNN to identify supraventricular tachycardia mechanisms by 12-lead electrocardiogram in the test dataset

		Experienced cardiac electrophysiologists (%)				CNN (%)				
		Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV	F1
AVNRT	Adjudicator 1	75	64.0	73.8	68.1	91.7*	60.1†	46.6	95.0	61.8
	Adjudicator 2	56.7	56.0	77.3	50.0					
	Average	65.9*	60.0†	75.6	59.1					
AVRT	Adjudicator 1	62.2	74.0	63.9	75.0	78.4*	63.5†	30.5	93.5	44.0
	Adjudicator 2	64.9	52.1	50.0	73.1					
	Average	63.6*	63.05†	57.0	74.1					
AT	Adjudicator 1	69.2	70.1	81.8	70.1	61.5*	65.4†	10.1	96.4	17.4
	Adjudicator 2	30.8	59.8	50	63.0					
	Average	50.0*	65.0†	66	66.6					

AT = atrial tachycardia; AVNRT = atrioventricular nodal re-entrant tachycardia; AVRT = atrioventricular reciprocating tachycardia; CNN = convolutional neural network; NPV = negative predictive value; PPV = positive predictive value.

*The average electrophysiologist sensitivity and the CNN sensitivity, at the fixed specificity.

†The average electrophysiologist specificity and the CNN specificity approximately fixed to this value for each class.

Table 4 Performance of CNN among subsets of long and short RP tachycardia in the test dataset

	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	F1 (95% CI)
Long RP tachycardia CNN						
AVNRT	0.741 (0.514–0.923)	40.0% (0.000–0.779)	89.3% (0.778–0.967)	47.7% (0.000–0.881)	87.5% (0.714–0.963)	42.9% (0.000–0.778)
AT	0.727 (0.491–0.850)	64.3% (0.375–0.848)	83.3% (0.654–0.957)	63.6% (0.363–0.858)	83.3% (0.684–0.945)	63.6% (0.429–0.800)
Short RP tachycardia CNN						
AVNRT	0.930 (0.893–0.954)	86.7% (0.805–0.928)	88.1% (0.830–0.940)	75.4% (0.643–0.846)	94.1% (0.910–0.970)	80.9% (0.722–0.863)
AVRT	0.887 (0.841–0.931)	51.5% (0.373–0.675)	96.2% (0.936–0.986)	76.5% (0.624–0.895)	89.1% (0.848–0.933)	61.2% (0.454–0.737)

Due to the very small sample size in respective substrata, AVRT in long RP tachycardia strata and AT in short RP tachycardia strata were excluded from the analysis.
AT = atrial tachycardia; AUC = area under the receiver-operating characteristic curve; AVNRT = atrioventricular nodal re-entrant tachycardia; AVRT = atrioventricular reciprocating tachycardia; CI = confidence interval; CNN = convolutional neural network; NPV = negative predictive value; PPV = positive predictive value.

P-wave to the QRS during tachycardia. However, most studies restrict the analysis to short RP tachycardia alone, and few presently available approaches allow comprehensive differentiation of SVT mechanisms across both long and short RP tachycardias.[4–12] Previous studies proposed algorithms that have been shown to diagnose approximately 70% to 90% of cases of AVRT and AVNRT in settings of short RP tachcyardia.[4–12] Because long RP tachycardias are difficult to characterize using standard clinical approaches with surface 12-lead ECG alone, they often require a comprehensive and invasive cardiac electrophysiologic study to confirm the underlying mechanism, which informs the type of ablation required. The present study demonstrates real-world application of a CNN artificial intelligence algorithm to adjudicate all cases of SVT forms, including in long RP tachycardia using surface 12-lead ECGs.

Clinical benefit of utilizing the CNN for 12-lead ECGs

Due to its reproducibility, widespread availability, and low cost, the surface 12-lead ECG is very well positioned to assist with providing a noninvasive SVT diagnosis. According to the current clinical paradigm, SVTs are divided into short RP tachycardias in which the P-wave is situated before the half way point of the RR interval. This configuration can be further divided into those with typical AVNRT and AVRT supported by retrograde conduction over a robust accessory pathway. Several studies have evaluated the value of the 12-lead ECG parameters for the differentiation of those 2 mechanisms, for example, the presence of pseudo r' waves in V1, or pseudo s waves in the inferior leads[4–8] and notches in the aVL lead[9] favor the diagnosis of typical AVNRT, while visible retrograde P waves or an RP interval of >100 ms,[4–8] QRS alternans,[4,6,8,10] marked repolarization change,[5,11] and ST-segment elevation in aVR[12] favor the diagnosis of AVRT. They demonstrated that those algorithms were able to accurately diagnose approximately 70% to 90% of cases of AVRT and AVNRT, which were compatible with our adjudication results in the short RP tachycardia subset (Table 5).[4–12] However, from a practical standpoint, the important limitation of those studies was that the results were not representative of all cases of SVT. For example, in the latter analysis of the SVT mechanism, the so-called long RP tachycardias as well as those with bundle branch block during tachycardia are likely to be more challenging to discern the mechanisms, and were not included in the evaluation. The long RP tachycardias raise the possibility of AT, atypical AVNRT, or AVRT using a decremental accessory pathway, making ECG characterization difficult.

Our results demonstrate that CNN analysis of 12-lead ECGs may help to augment clinical assessment for AVNRT and AVRT SVT mechanisms, wherein the CNN outperformed electrophysiology experts, and also for AT, which showed high specificity, though sensitivity was lower. Overall, the CNN's high specificities even among long RP

Table 5 Performance of the experienced cardiac electrophysiologists vs CNN among subsets of long RP and short RP tachycardia

	Experienced cardiac electrophysiologists (%)				CNN (%)				
	Sensitivity	Specificity	PPV	NPV	Sensitivity	Specificity	PPV	NPV	F1
Long RP tachycardia rowhead									
AVNRT	25.9	78.1	30.4	74.0	100	78.1	8.5	100	15.7
AT	54.9	75.9	84.8	40.7	100	75.0	15.8	100	27.3
Short RP tachycardia rowhead									
AVNRT	74.5	74.5	83.3	63.1	91.5	75.3	53.1	96.7	67.2
AVRT	72.3	76.4	64.2	82.6	91.4	76.5	45.1	97.7	60.4

Due to the very small sample size in respective substrata, AVRT in long RP tachycardia strata and AT in short RP tachycardia strata were excluded from the analysis.

AT = atrial tachycardia; AVNRT = atrioventricular nodal re-entrant tachycardia; AVRT = atrioventricular reciprocating tachycardia; CNN = convolutional neural network; NPV = negative predictive value; PPV = positive predictive value.

tachycardia across all SVT mechanisms may help clinicians to rule out cases with low pretest probability for specific SVT mechanisms.

Future directions for developing CNNs

This study is the first to our knowledge to demonstrate strong CNN performance for SVT mechanism as adjudicated by a gold-standard electrophysiology study, and it corroborates recent studies supporting strong CNN performance for various ECG-based diagnoses. Our CNN algorithm provides clinically relevant future directions to improve the performance of CNNs for SVT diagnoses using a 12-lead ECG alone. First, because CNNs incrementally improve as more training data are made available,[31,32] greater amounts of training data (ideally similarly adjudicated by gold-standard electrophysiology studies) would likely improve CNN performance for all 3 SVT mechanisms and possibly allow expansion to other tachyarrhythmic mechanisms to achieve the fullest potential of CNN-augmented ECG diagnosis. For example, recent work from Luongo and colleagues[14] demonstrated that a machine learning classifier relying only on noninvasive signals allowed identifying the origin of atrial flutter. Second, a well-trained CNN has the ability to identify novel ECG associations beyond what electrophysiology doctors can perform with ECGs,[15,17–19] raising the possibility of expanding the diagnostic utility of ECGs beyond their present scope. Ultimately, ECG analysis systems that consistently achieve expert-level performance may potentially allow for changing the clinical workflow, either by providing accurate triage and screening for certain diagnoses or by providing real-time feedback to clinicians reading ECGs in the electrophysiology lab or inpatient or outpatient clinics.

Limitations

Several limitations are evident in this study. We did not perform an external validation, raising the possibility of a lower CNN performance with data from another institution on account of overfitting or the systematic differences in ECGs. Also, the adjudication results from the 2 electrophysiologists display variability, reflecting the real-world difficulties in precisely diagnosing SVT mechanisms using 12-lead ECG alone even among experienced electrophysiology clinicians. A relatively large number of patients had to be excluded for lack of induction of sustained arrhythmia or technical problems as well as for having rare arrhythmias that were felt to be important to remove confounders only applicable to a very small part of the SVT population. For example, the clinical finding of SVT with 2:1 AV block allows for ready exclusion of AVRT as an etiology. We thus excluded 2:1 AV block SVT from our population in order to allow for facile comparison between clinicians and the CNN, making the intended use of our SVT algorithm limited to those with 1:1 AV conduction. Training any algorithm on lower-prevalence diseases, such as AT here, presents challenges for training and evaluation, as low prevalence affects F1 score and PPV, making it challenging to compare between classes with variable prevalence. Future work may aim to improve this by increasing the dataset size for the rare classes, such as through cross-institutional collaborations. Therefore, larger datasets are required to train deep learning–based algorithms across the full range of SVT diagnoses. In addition, because stratification by long or short RP tachycardia decreased sample size, we opted to present results of our sensitivity analysis comparing the CNN with the electrophysiologists in these strata using all available adjudicated examples. The limitation in doing this is that the CNN performance in this analysis (Table 5) may be overly optimistic because some examples were drawn from patients in the training dataset. However, Table 4 shows CNN performance alone in the test dataset, and these results provide a foundation for future work with larger patient cohorts of long RP tachycardia.

Conclusion

In this study, 12-lead surface ECG data were used to train a CNN algorithm that achieved strong performance against gold-standard electrophysiology study–derived diagnosis and comparable performance to experienced cardiac

electrophysiology doctors. Larger datasets may be required to train deep learning–based algorithms across the full range of SVT diagnoses.

Funding Sources: Geoffrey H. Tison received support for this work from the National Institutes of HealthNational Institutes of Health/National Heart, Lung, and Blood Institute K23HL135274.

Disclosures: Geoffrey H. Tison has received research grants from Myokardia Inc, a fully owned subsidiary of Bristol Meyers Squibb, Janssen, and General Electric; and is an advisor to Viz.ai and Prolaio.

Authorship: All authors attest they meet the current ICMJE criteria for authorship.

Patient Consent: All electrophysiology studies were initially performed with a signed informed consent; the need for informed consent for the research analysis was waived by the institutional review board in the setting of anonymized retrospective record review.

Ethics Statement: This study was conducted according to the principles of the Declaration of Helsinki, and the study protocol was approved by the University of California San Francisco Institutional Review Board.

References

1. Ávila P, Calvo D, Tamargo M, et al. Association of age with clinical features and ablation outcomes of paroxysmal supraventricular tachycardias. Heart 2022; 108:1107–1113.
2. Walsh MA, Gonzalez CM, Uzun OJ, et al. Outcomes from pediatric ablation: a review of 20 years of national data. J Am Coll Cardiol EP 2021;7:1358–1365.
3. Porter MJ, Morton JB, Denman R, et al. Influence of age and gender on the mechanism of supraventricular tachycardia. Heart Rhythm 2004;1:393–396.
4. Kalbfleisch SJ, el-Atassi R, Calkins H, Langberg JJ, Morady F. Differentiation of paroxysmal narrow QRS complex tachycardias using the 12-lead electrocardiogram. J Am Coll Cardiol 1993;21:85–89.
5. Erdinler I, Okmen E, Oguz E, Akyol A, Gurkan K, Ulufer T. Differentiation of narrow QRS complex tachycardia types using the 12-lead electrocardiogram. Ann Noninvasive Electrocardiol 2002;7:120–126.
6. González-Torrecilla E, Almendral J, Arenal A, Atienza F, del Castillo S, Fernández-Avilés F. Independent predictive accuracy of classical electrocardiographic criteria in the diagnosis of paroxysmal atrioventricular reciprocating tachycardias in patients without pre-excitation. Europace 2008;10:624–628.
7. Tai CT, Chen SA, Chiang CE, et al. A new electrocardiographic algorithm using retrograde P waves for differentiating atrioventricular node reentrant tachycardia from atrioventricular reciprocating tachycardia mediated by concealed accessory pathway. J Am Coll Cardiol 1997;29:394–402.
8. Letsas KP, Weber R, Siklody CH, et al. Electrocardiographic differentiation of common type atrioventricular nodal reentrant tachycardia from atrioventricular reciprocating tachycardia via a concealed accessory pathway. Acta Cardiol 2010;65:171–176.
9. Di Toro D, Hadid C, López C, Fuselli J, Luis V, Labadet C. Utility of the aVL lead in the electrocardiographic diagnosis of atrioventricular node re-entrant tachycardia. Europace 2009;11:944–948.
10. Green M, Heddle B, Dassen W, et al. Value of QRS alteration in determining the site of origin of narrow QRS supraventricular tachycardia. Circulation 1983; 68:368–373.
11. Jaeggi ET, Gilljam T, Bauersfeld U, Chiu C, Gow R. Electrocardiographic differentiation of typical atrioventricular node reentrant tachycardia from atrioventricular reciprocating tachycardia mediated by concealed accessory pathway in children. Am J Cardiol 2003;91:1084–1089.
12. Ho YL, Lin LY, Lin JL, Chen MF, Chen WJ, Lee YT. Usefulness of ST-segment elevation in lead aVR during tachycardia for determining the mechanism of narrow QRS complex tachycardia. Am J Cardiol 2003;92:1424–1428.
13. Stevens LM, Mortazavi BJ, Deo RC, Curtis L, Kao DP. Recommendations for reporting machine learning analyses in clinical research. Circ Cardiovasc Qual Outcomes 2020;13:e006556.
14. Luongo G, Vacanti G, Nitzke V, et al. Hybrid machine learning to localize atrial flutter substrates using the surface 12-lead electrocardiogram. Europace 2022; 24:1186–1194.
15. Hughes JW, Olgin JE, Avram R, et al. Performance of a convolutional neural network and explainability technique for 12-lead electrocardiogram interpretation. JAMA Cardiol 2021;6:1285–1295.
16. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med 2019;25:65–69.
17. Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet 2019; 394:861–867.
18. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat Med 2019;25:70–74.
19. Tison GH, Zhang J, Delling FN, Deo RC. Automated and interpretable patient ECG profiles for disease detection, tracking, and discovery. Circ Cardiovasc Qual Outcomes 2019;12:e005289.
20. Katritsis DG, Josephson ME. Differential diagnosis of regular, narrow-QRS tachycardias. Heart Rhythm 2015;12:1667–1676.
21. Michaud GF, Tada H, Chough S, et al. Differentiation of atypical atrioventricular node re-entrant tachycardia from orthodromic reciprocating tachycardia using a septal accessory pathway by the response to ventricular pacing. J Am Coll Cardiol 2001;38:1163–1167.
22. Padanilam BJ, Manfredi JA, Steinberg LA, Olson JA, Fogel RI, Prystowsky EN. Differentiating junctional tachycardia and atrioventricular node re-entry tachycardia based on response to atrial extrastimulus pacing. J Am Coll Cardiol 2008;52:1711–1717.
23. Ho RT, Frisch DR, Pavri BB, Levi SA, Greenspon AJ. Electrophysiological features differentiating the atypical atrioventricular node-dependent long RP supraventricular tachycardias. Circ Arrhythm Electrophysiol 2013;6:597–605.
24. Yeo IK, Johnson RA. A new family of power transformations to improve normality or symmetry. Biometrika 2000;87:954–959.
25. Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006;313:504–507.
26. Nair V, Hinton G. Rectified linear units improve Restricted Boltzmann machines. Proceedings of International Conference on Machine Learning 2010;27:807–814.
27. Ioffe S, Szegedy C. Batch normalization : accelerating deep network training by reducing internal covariate shift. 2015. Proceedings of the 32nd International Conference on Machine Learning 2015;37:448–456.
28. Kingma DP, Ba JL. Adam: A method for stochastic optimization. Presented at: 3rd International Conference on Learning Representations, ICLR 2015; May 7–9, 2015; San Diego, CA.
29. Chollet F. Keras. GitHub. 2015. Available at: https://github.com/fchollet/keras. Accessed July 24, 2023.
30. Rossum GV. Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam. May 1995. Available at: https://ir.cwi.nl/pub/5007/05007D.pdf. Accessed July 24, 2023.
31. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402–2410.
32. Sun C, Shrivastava A, Singh S, Gupta A. Revisiting unreasonable effectiveness of data in deep learning era. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. 2017. p. 843–852.