Journal of Electrocardiology 57 (2019) 39–43

Contents lists available at ScienceDirect

Journal of Electrocardiology

journal homepage: www.jecgonline.com

Data driven feature selection and machine learning to detect misplaced V1 and V2 chest electrodes when recording the 12 lead electrocardiogram

Khaled Rjoob, MSc [a,*], Raymond Bond, PhD [a], Dewar Finlay, PhD [a], Victoria McGilligan, PhD [b], Stephen J. Leslie, PhD [c], Aleeha Iftikhar, MSc [a], Daniel Guldenring, PhD [d], Ali Rababah, MSc [a], Charles Knoery, MSc [c], Anne McShane, MSc [e], Aaron Peace, PhD [f]

[a] Faculty of Computing, Engineering & Built Environment, Ulster University, Northern Ireland, UK
[b] Faculty of Life & Health Sciences, Centre for Personalised Medicine, Ulster University, Northern Ireland, UK
[c] Department of Diabetes & Cardiovascular Science, University of the Highlands and Islands, Centre for Health Science, Inverness, UK
[d] HTW Berlin, Wilhelminenhofstr. 75A, 12459 Berlin, Germany
[e] Emergency Department, Letterkenny University Hospital, Donegal, Ireland
[f] Western Health and Social Care Trust, C-TRIC, Ulster University, Northern Ireland, UK

ARTICLE INFO

Available online xxxx

Keywords:
Machine learning
Feature extraction
Body surface potential maps
Lead misplacement
Electrode misplacement
Chest leads

ABSTRACT

Background: Electrocardiogram (ECG) lead misplacement can adversely affect ECG diagnosis and subsequent clinical decisions. V1 and V2 are commonly placed superior of their correct position. The aim of the current study was to use machine learning approaches to detect V1 and V2 lead misplacement to enhance ECG data quality.

Method: ECGs for 453 patients, (normal n = 151, Left Ventricular Hypertrophy (LVH) n = 151, Myocardial Infarction n = 151) were extracted from body surface potential maps. These were used to extract both the correct and incorrectly placed V1 and V2 leads. The prevalence for correct and incorrect leads were 50%. Sixteen features were extracted in three different domains: time-based, statistical and time-frequency features using a wavelet transform. A hybrid feature selection approach was applied to select an optimal set of features. To ensure optimal model selection, five classifiers were used and compared. The aforementioned feature selection approach and classifiers were applied for V1 and V2 misplacement in three different positions: first, second and third intercostal spaces (ICS).

Results: The accuracy for V1 misplacement detection was 93.9%, 89.3%, 72.8% in the first, second and third ICS respectively. In V2, the accuracy was 93.6%, 86.6% and 68.1% in the first, second and third ICS respectively. There is a noticeable decline in accuracy when detecting misplacement in the third ICS which is expected.

Introduction

The electrocardiogram (ECG) comprises of several processes including patient preparation and electrode placement [1]. Lead misplacement has a large effect on ECG morphology which can lead to an effect on the ECG interpretation and diagnosis [2,3]. It can occur between 40% and 60% of the time and 36% of chest electrodes are placed outside a radius of 1.25 in. from the correct positions [4] [5]. It has also been reported that only 50% of nurses and 20% of cardiologists can correctly place the electrodes V1 and V2 in their correct position [6]. Changes in morphological features of the ECG are notable with misplacements with distances greater than 2 cm to the standard electrode position [2]. A common mistake in chest lead misplacement is misplacing leads V1 and V2 too high, which can also offset the remaining precordial electrodes (V3 to V6) [7]. V5 and V6 are also commonly placed in the fifth intercostal space (ICS) and are not in the recommended position (V5 should be placed in the fifth ICS and left anterior axillary line, and v6 should be placed on the left mid-axillary line at same horizontal level as V5) [8] [9] [10]. The analysis of ECG signals that were recorded from vertically misplaced V1 and V2 leads could result in a false diagnosis of Brugada syndrome and anterior STEMI, which can lead to an incorrect diagnosis in 17–24% of patients [11–13]; because lead misplacement can conceal or mimic another cardiac disease [14–17]. Several techniques have been suggested and used to control the deviation in precordial leads positioning. One of these methods involved a sliding ruler to facilitate proper electrode placement in the correct positions while, another technique was placing stickers on a diagram of

* Corresponding author.
E-mail address: rjoob-k@ulster.ac.uk (K. Rjoob).

https://doi.org/10.1016/j.jelectrocard.2019.08.017
0022-0736/

torso to improve electrode placement and another simple technique was electrode belts [18,19]. Unfortunately, clinicians have not yet widely adopted these techniques; probably due to their cost or insufficient evidence of their efficiency [15]. In previous studies, ML has been used to detect limb lead misplacement with promising results [20], but has not be tested in the setting of misplaced chest electrodes. We therefore aimed to use ML to detect chest electrode misplacement.

Material and methods

Data collection

A high-resolution body surface potential maps (BSPM) were used to extract ECG signals for V1 and V2 leads. Each BSPM consists of 117 nodes which represent 117 ECG leads. The dataset used has been recorded during the 1980s by Fred Kornreich and is known as the Kornreich dataset [21–23]. The unique nature of the dataset has seen it widely utilised in studies where large volumes of simultaneously recorded leads are desired. The 117 nodes in each BSPM have been multiplied with transformation matrix to get 352 nodes that provide greater resolution. The ECG dataset is comprised of three different subjects including 151 normal, 151 Myocardial Infarction (MI) and 151 Left Ventricular Hypertrophy (LVH) patients. All patients with MI had a typical history of prolonged, ischemic-type cardiac pain and characteristic changes in enzyme levels. In most cases, the diagnosis was further substantiated by coronary angiography and ventriculography, echocardiography, or nuclear angiograms during which multigated, technetium99m-labeled blood pool imaging was performed. We extracted ECG signals for V1 and V2 in their correct and incorrect positions in three ICS (1st, 2nd and 3rd). Class labels were balanced where 50% of cases are correct and 50% are incorrect ECGs to avoid bias. Nodes numbered 169 and 171 in the green color (refer to Fig. 1) represent both leads V1 and V2 respectively in their correct positions.

Nodes 126 and 128 in the blue color represent the misplaced leads V1 and V2 in the third ICS, nodes 83 and 85 in blue color represent V1 and V2 misplacement in the second ICS and node 43 and 45 in blue color represent V1 and V2 misplacement in the first ICS. The ECG signals were recoded simultaneously from each patient for leads V1 and V2.

Feature extraction

ECG features were extracted in three different domains: 1) Time domain: we considered P wave amplitude (regardless to monophasic and biphasic P waves), PR interval, QRS onset value, R-wave peak amplitude, offset of the QRS and S wave amplitude, 2) Statistical domain: mean, standard deviation, skewness, and kurtosis of the ECG signal, and 3) Time-frequency features: discrete wavelet transform (DWT) was applied using 4 levels and symlets wavelet mother function. We retrieved four detailed coefficients (D1, D2, D3, D4). We considered the first maximum three values of D3 as features. Table 1 describes feature ID, domain and description.

Feature selection

A selection technique that is used to provide a subset of features that provides a good classification result. The Filter method and the wrapper method have been applied sequentially. Firstly, Filter method including mutual information feature selection (MIFS) using Eq. (1), maximum relevance minimum redundancy (MRMR) using Eq. (1), joint mutual information JMI using Eq. (1), Entropy using Eq. (2) and Relief using Eq. (3) has been used to rank feature from the most important feature to least important feature. Secondly, wrapper method using a backwards elimination algorithm has been applied on ranked features to find an optimal set of features as inputs to the ML classifiers. The backwards elimination algorithm starts with all 16 features and removes features, one by one, the point at which removing a feature does not improve performance.

$$f_t = \arg\max I(x_i;y) - \left[\alpha \sum_{k=1}^{t-1} I(x_{fk};x_i) - \beta \sum_{k=1}^{t-1} I(x_{fk};x_i|y)\right] \quad (1)$$

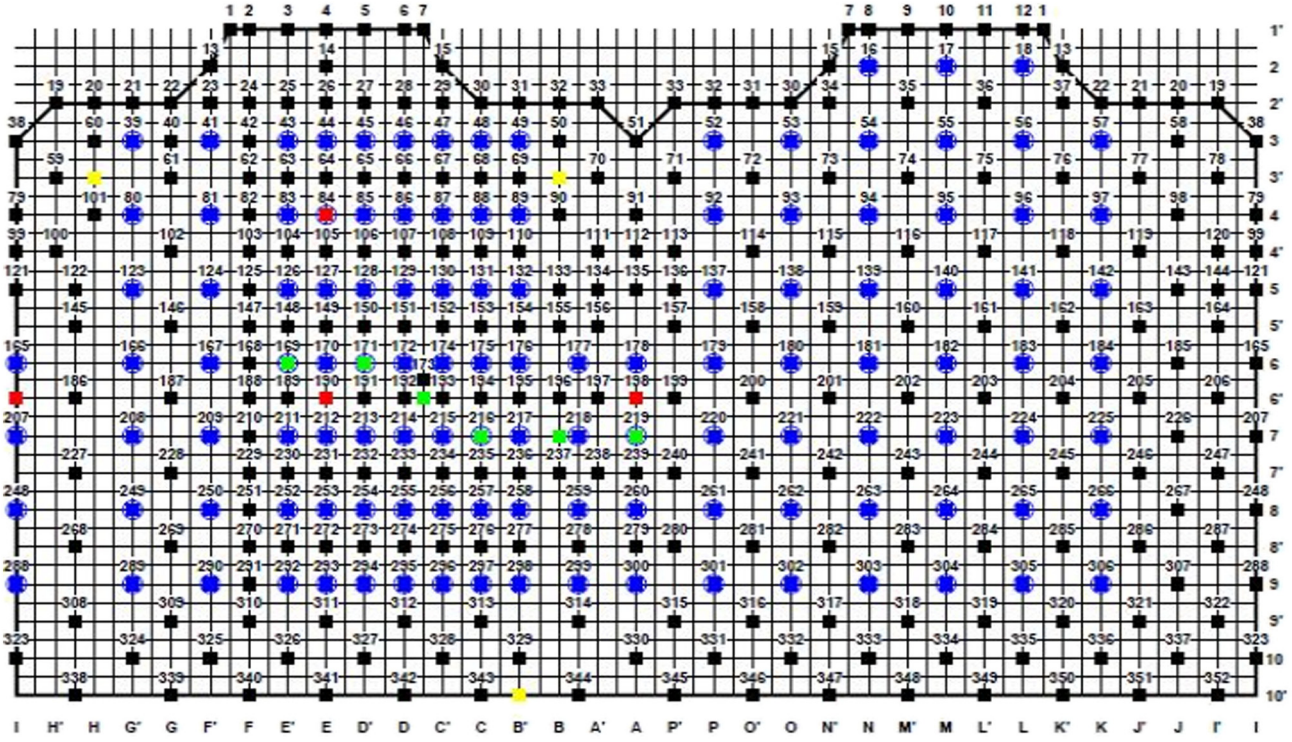

Fig. 1. Dalhousie torso with 352 nodes.

Table 1
Feature ID, domain and description.

Feature ID	Domain	Feature description
1	Time domain	P wave amplitude
2	Time domain	PR interval
3	Time domain	QRS beginning value/P wave amplitude
4	Time domain	P wave amplitude/QRS beginning
5	Time domain	R wave/end of QRS
6	Time domain	S wave amplitude
7	Time domain	End of QRS/R wave
8	Statistical domain	Mean of ECG signal
9	Statistical domain	Variance of ECG signal
10	Statistical domain	Standard deviation of ECG signal
11	Statistical domain	Skewness of ECG signal
12	Statistical domain	Kurtosis of ECG signal
13	Time-Frequency domain	First maximum value of D3/Third maximum value of D3
14	Time-Frequency domain	First maximum value of D3/Second maximum value of D3
15	Time-Frequency domain	Second maximum value of D3/Third maximum value of D3
16	Time-Frequency domain	First maximum value of D3/First maximum value of D3

where x represents features and y represents labels

$$\text{In JMI}: \alpha = \frac{1}{t-1} \text{ and } \beta = \frac{1}{t-1}$$

$$\text{In MIFS}: \alpha = 0 \text{ and } \beta = 0$$

$$\text{In MRMR}: \alpha = \frac{1}{t-1} \text{ and } \beta = 0$$

$$H(X) = -\sum p(x)\ \log_2 p(x) \tag{2}$$

where x represents features

$$W_i = W_i - (x_i - nearHit_i)^2 + (x_i - nearMiss_i)^2 \tag{3}$$

w is the weighted vector initialised with zeros, where x represents features, nearHit is the closest same class instance and nearMiss is the closest different class instance.

Classification

V1 and V2 placement were classified separately to detect misplacement. Five machine learning classifiers have been used to get the best possible accuracy including: 1) Fine tree (FTree), 2) Coarse tree (CTree), 3) Logistic regression (LOGT), 4) Linear support vector machine (LSVM) and 5) Quadratic support vector machine (QSVM). Machine learning classifiers performance were evaluated using accuracy (Eq. (4)), sensitivity (Eq. (5)) and specificity (Eq. (6)). A true positive test result was defined as incorrect electrode placement correctly identified as incorrect and a true negative was defined as correct electrode placement correctly identified as correct.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

Results

The best accuracy for V1 misplacement detection as shown in Table 2 was 93.9%, 89.3%, 72.8%, sensitivity was 96.6%, 89.9%, 71.1%

Table 2
Lead misplacement detection accuracy of V1 and V2 using ML in each ICS.

	V1 1st intercostal space					V2 1st intercostal space				
	JMI[15f]	MIFS[16f]	MRMR[13f]	RELIEF[16f]	ENTR[15f]	JMI[15f]	MIFS[13f]	MRMR[13f]	RELIEF[16f]	ENTR[15f]
FTREE	89.9%	89.9%	91.6%	87.2%	89.3%	89.3%	87.2%	87.9%	91.3%	90.9%
CTREE	88.3%	84.2%	85.9%	86.6%	89.3%	86.2%	87.2%	86.2%	88.3%	86.9%
LOGT	87.9%	84.9%	85.2%	83.9%	83.2%	91.3%	89.3%	87.6%	85.6%	85.9%
LSVM	87.6%	82.9%	82.9%	82.6%	83.6%	88.6%	88.6%	84.9%	86.9%	84.9%
QSVM	**93.9%**	90.3%	87.9%	88.9%	91.6%	**93.6%**	**93.6%**	89.3%	87.2%	92.3%
	V1 2nd intercostal space					V2 2nd intercostal space				
	JMI[15f]	MIFS[16f]	MRMR[16f]	RELIEF[15f]	ENTR[16f]	JMI[15f]	MIFS[16f]	MRMR[12f]	RELIEF[16f]	ENTR[15f]
FTREE	85.2%	88.6%	82.6%	84.9%	84.2%	78.2%	**86.6%**	84.6%	79.5%	83.6%
CTREE	83.6%	82.2%	81.9%	85.9%	86.2%	75.8%	79.9%	79.2%	79.5%	79.5%
LOGT	82.9%	82.9%	81.9%	82.6%	81.5%	76.2%	79.5%	79.5%	75.5%	81.9%
LSVM	84.9%	84.2%	83.6%	84.6%	82.6%	77.5%	80.9%	83.2%	76.2%	80.2%
QSVM	85.6%	86.6%	86.2%	86.9%	**89.3%**	80.9%	84.2%	81.9%	77.5%	84.2%
	V1 3rd intercostal space					V2 3rd intercostal space				
	JMI[15f]	MIFS[15f]	MRMR[16f]	RELIEF[16f]	ENTR[15f]	JMI[15f]	MIFS[15f]	MRMR[12f]	RELIEF[15f]	ENTR[15f]
FTREE	63.4%	60.1%	69.1%	65.8%	61.7%	58.7%	60.1%	59.1%	60.4%	57.7%
CTREE	63.4%	69.5%	68.5%	71.1%	68.5%	58.7%	59.1%	61.7%	63.1%	57.0%
LOGT	70.1%	**72.8%**	67.4%	67.8%	70.8%	67.1%	66.4%	64.8%	68.5%	62.1%
LSVM	70.8%	71.1%	66.8%	70.8%	70.1%	64.8%	64.4%	61.4%	63.1%	62.4%
QSVM	65.4%	68.1%	69.5%	68.5%	65.1%	61.1%	**68.1%**	65.1%	65.4%	61.1%

Bold represent the best accuarcy in each intercostal space in V1 and V2.
FTREE: fine tree, CTREE: coarse tree, LOGT: logistic regression, LSVM: linear support vector machine, QSVM: quadratic support vector machine, JMI: joint mutual information, MIFS: mutual information feature selection, MRMR: maximum relevance minimum redundancy, ENTR: entropy. Superscripts such as MIFS[13f] represent the best number of features to provide a good accuracy in feature selection algorithm.

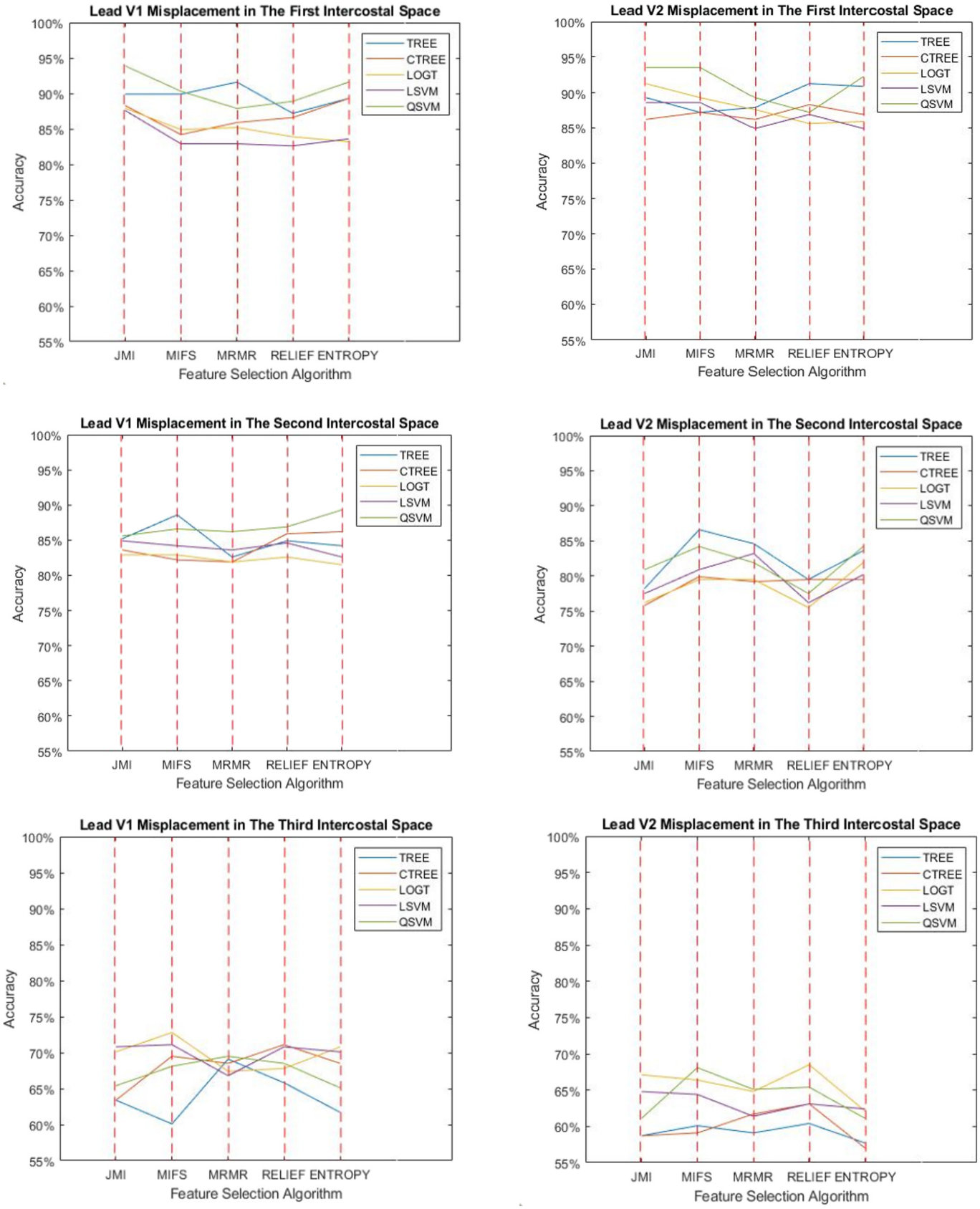

Fig. 2. Accuracy in each ICS using five ML classifiers. Each subplot represents accuracy in each combination between ML classifiers and feature selection algorithms in 1st, 2nd and 3rd ICS.

and specificity was 91.2%, 88.5%, 74.5% in the first, second and third ICS respectively. In V2, the best accuracy for lead misplacement detection was 93.6%, 86.6% and 68.1%, sensitivity was 95.3%, 86.6%, 67.1% and specificity was 91.9%, 86.6%, 69.1% in the first, second and third ICS respectively. Fig. 2 shows accuracy for each classifier combined with each feature selection algorithm in three ICS and QSVM provides the best performance for detecting misplacement.

In the first ICS, in V1 the best classifier was QSVM and the best feature selection algorithm was JMI and in V2 the best classifier was QSVM and the best feature selection algorithm was JMI and MIFS. In the second ICS, the best classifier was QSVM combined with Entropy for V1 and FTREE combined with MIFS for V2, while in the third ICS, the best performance was for LOGT combined with MIFS for V1 and QSVM combined with MIFS for V2.

Discussion

In this study, a new method was applied and achieved a high accuracy to detect chest lead misplacement, particularly in the first and second ICS, while, there is a noticeable decline in accuracy when there is lead misplacement in the third ICS. This is theoretically expected, since the ECG features will be more similar to features recorded in the correct 4th ICS. QSVM and FTREE have the best performance (from 86.6% to 93.9%) in the first and second ICS using JMI, MIFS and entropy as a feature selection algorithm. This study used the same morphological features in previous researches, but with using some different classifiers like QSVM and by adding new features, like time-frequency feature and statistical features. The developed algorithm could help clinicians to reduce the number of incorrectly performed ECGs and reduce time and cost of performing a new ECG and as a consequence, this could help physicians to get an accurate and quick diagnosis using correct recorded ECG. Future research will therefore aim to improve detecting of chest lead misplacement accuracy in the third ICS using alternative techniques or by showing the probability of that lead being misplaced to enable clinicians to decide if they need to record the ECG again. The inclusion of new features such as relative features that for example incorporates similarities and ratios between features extracted from the two leads (e.g. correlation between V1 and V2 or the ratio between amplitudes such as the R peaks) will also be investigated. Further experimentation with deep learning should also be carried out to detect lead misplacement and compare the machine learning models with the performance of criteria such as rSr′ for classifying V1 misplacement [24]. Alternatively, based on these results, a medical device company may only want to deploy an algorithm to detect leads misplaced in the second intercostal space until the accuracy of detecting leads placed in the third intercostal space is improved. In this way, ECG recording quality can be iteratively improved. A comparison of the performance of these models with the performance of healthcare professionals in classifying V1 and V2 lead misplacement will lead to further insights and possible improvements.

Conclusion

In this work, QSVM and FTree provide the best performance for detecting chest lead misplacement in the 1st and 2nd ICS QSVM, while LOGT and QSVM have the best performance in the 3rd ICS. Based on feature selection algorithm, in the 1st ICS the best feature selection algorithm was JMI. In the 2nd ICS, the best algorithm was entropy and MIFS, while in the 3rd ICS MIFS performs better than the other feature selection algorithm. Based on results, a derivation study should be carried out using a broader dataset to check if the invented algorithm can improve ECG data quality and decision making in cardiac care. And to prospect the effects of type 1 and type 2 errors.

Acknowledgements

We thank ISCE conference participants for comments that improved the manuscript and provided insight that assisted the research.

Author contributors

All authors were responsible for study conception. All were responsible for design, analysis and interpretation of results. All were responsible for revising the manuscript.

Declaration of competing interest

None.

Funding

This work is supported by the European Union's INTERREG VA programme, managed by the Special EU Programmes Body (SEUPB). The views and opinions expressed in this study do not necessarily reflect those of the European Commission or the Special EU Programmes Body (SEUPB).

References

[1] The Society for Cardiological Science and Technology. Recording a Standard 12-Lead Electrocardiogram an Approved Methodology. British Cardiovascular society; 2013.
[2] Schijvenaars BJ, et al. Effect of electrode positioning on ECG interpretation by computer. J Electrocardiol 1997;30:247–56.
[3] Richard L, et al. ECG Lead misplacement: a brief review of limb lead misplacement. African Journal of Emergency Medicine 2014;4:130–9.
[4] Wenger W, Kligfield P. Variability of precordial electrode placement during routine electrocardiography. J Electrocardiol 1996;29:179–84.
[5] Bupp JE, et al. Placement of cardiac electrodes: written, simulated, and actual accuracy. Crit Care 1997;6:457–62.
[6] Khunti K. Accurate interpretation of the 12-lead ECG electrode placement: a systematic review. Health Educ J 2014;73:610–23.
[7] Rajaganeshan R, et al. Accuracy in ECG lead placement among technicians, nurses, general physicians and cardiologists. Int J Clin Pract 2008;62:65–70.
[8] Kligfield P, et al. Recommendations for the standardization and interpretation of the electrocardiogram. Part I: the electrocardiogram and its technology. A scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology, the American College of Cardiology Foundation; and the Heart Rhythm Society. Heart Rhythm 2007;4:394–412.
[9] Michał K, et al. The effect of precordial lead displacement on ECG morphology. Med Biol Eng Comput 2014;52:109–19.
[10] Kerwin AJ, et al. A method for the accurate placement of chest electrodes in the taking of serial electrocardiographic tracings. Can Med Assoc J 1960;82:258–61.
[11] Lehmann MH, et al. Brugada syndrome: diagnostic pitfalls. J Emerg Med 2009;37:79–81.
[12] Bond R, Finlay D, et al. The effects of electrode misplacement on clinicians' interpretation of the standard 12-lead electrocardiogram. Eur J Intern Med 2012;23(7):610–5.
[13] Bond R, et al. Human factors analysis of the CardioQuick Patch: a novel engineering solution to the problem of electrode misplacement during 12-lead electrocardiogram acquisition. J Electrocardiol 2016;49:911–8.
[14] Rudiger A, et al. Influence of electrode misplacement on the electrocardiographic signs of inferior myocardial ischaemia. Am J Emerg Med 2003;21:574–81.
[15] Chanarin N, et al. Pseudo reinfarction: a consequence of electrocardiogram lead transposition following myocardial infarction. Clin Cardiol 1990;13:228–9.
[16] Jowett NI, et al. Modified electrode placement must be recorded when performing 12-lead electrocardiograms. Postgrad Med J 2005;81:122–5.
[17] Brooks Walsh MD. Misplacing V1 and V2 can have clinical consequences. J Electrocardiol 2018;36:865–70.
[18] Herman MV, et al. Variability of electrocardiographic precordial lead placement: a method to improve accuracy and reliability. Clin Cardiol 1991;14:469–76.
[19] Bupp JE, Dinger M, et al. Placement of cardiac electrodes: written, simulated and actual accuracy. Am J Crit Care 1997;6:457–62.
[20] Richard E, et al. Detecting ECG limb lead-wire interchanges involving the right leg lead-wire. Computing in Cardiology 2017;44.
[21] Kornreich F, Montague TJ, Rautaharju PM. Identification of first acute Q wave and non-Q wave myocardial infarction by multivariate analysis of body surface potential maps. Circulation 1991;84(6):2442–53 Dec.
[22] Montague TJ, Smith ER, Cameron DA, Rautaharju PM, Klassen GA, Felmington CS, et al. Isointegral analysis of body surface maps: surface distribution and temporal variability in normal subjects. Circulation 1981;63(5):1166–72 May.
[23] Kornreich F, Lux RL, MacLeod RS. Map representation and diagnostic performance of the standard 12-lead ECG. J Electrocardiol 1995;28:121–3 Jan 1.
[24] Adrian B, et al. Differential diagnosis of rSr′ pattern in leads V1-V2. Comprehensive review and proposed algorithm. Ann Noninvasive Electrocardiol 2015;20:7–17.