2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Repeated Structuring & Learning Procedure for Detection of Myocardial Ischemia: a Robustness Analysis

Agnese Sbrollini, *Member, IEEE*, Ilaria Marcantoni, *Member, IEEE*, Micaela Morettini, *Member, IEEE*, Cees A. Swenne and Laura Burattini, *Member, IEEE*

Abstract— **Myocardial ischemia, consisting in a reduction of blood flow to the heart, may cause sudden cardiac death by myocardial infarction or trigger serious abnormal rhythms. Thus, its timely identification is crucial. The Repeated Structuring and Learning Procedure (RS&LP), an innovative constructive algorithm able to dynamically create neural networks (NN) alternating structuring and learning phases, was previously found potentially useful for myocardial ischemia detection. However, performance of created NN depends on three parameters, the values of which need to be set a priori by the user: maximal number of layers (NL), maximal number of initializations (NI) and maximal number of confirmations (NC). A robustness analysis of RS&LP to varying values of NL, NI and NC is fundamental for clinical applications concerning myocardial ischemia detection but was never performed before; thus, it was the aim the present study. Thirteen serial ECG features were extracted by pairs of ECGs belonging to 84 cases (patients with induced myocardial ischemia) and 398 controls (patients with no myocardial ischemia) and used as inputs to learn (50% of population) and test (50% of population) NNs with varying values of NL (1,2,3,4,10), NI (50,250,500,1000,1500) and NC (2,5,10,20,50). Performance of obtained NNs was compared in terms of area under the curve (AUC) of the receiver operating characteristics. Overall, 13 NNs were considered; 12 (92%) were characterized by AUC≥80% and 4 (31%) by AUC≥85%. Thus, RS&LP proved to be robust when creating NNs for detecting of myocardial ischemia.**

Clinical Relevance— **Availability of the Repeated Structuring & Learning Procedure to create reliable neural networks for timely detection of myocardial ischemia.**

I. Introduction

Ischemic heart disease is one of the main causes of death worldwide [1]. It consists in a reduction of blood flow to the heart, with consequent reduction of myocardial oxygen supply and cardiac pumping ability. The most critical consequence of ischemia is sudden cardiac death by myocardial infarction or triggering of serious abnormal rhythms. Thus, timely clinical decisions are essential in case of myocardial ischemia. In a patient with symptoms, an accurate triage already in the ambulance may allow timely action [2] to preserve cardiac functions as much as possible [3].

Clinically, diagnosis of ischemia and, in general of emerging pathologies, is usually performed by serial electrocardiography, which consists in the comparison of a newly acquired electrocardiogram (ECG) with a previously acquired ECG from the same subject [4]. Through serial ECG comparison, clinicians try to identify ECG modifications caused by a changed clinical status and not to cardiac physiological adaptations [5] and slightly different electrode positioning [6]. The emergency context of myocardial ischemia diagnosis and the complexity of serial ECG clinical interpretation make availability of automatic diagnostic support tools desirable. Among these, the ones based on artificial intelligence and, specifically, on machine learning approaches seem to be particularly promising.

Among the several studies for automatic detection of cardiac diseases [7–10], only a few of them deal with machine learning applications to serial electrocardiography [2, 11, 12]. Important but preliminary results were obtained by the new Repeated Structuring and Learning Procedure (RS&LP) [11], which was applied to create a neural networks (NN) able to detect myocardial ischemia on basis of thirteen serial ECG features provided as inputs. The most innovative aspects of RS&LP consist in its ability to dynamically create a NN by avoiding an initial definition of its structure and by performing an iterative learning, aspects that allowed RS&LP to reach better performance than NNs based on standard learning [11]. In its first use, RS&LP was implemented with the following three parameters a priori experimentally set by the user: maximal number of layers, maximal number of initializations and maximal number of confirmations. Since changes in the value of these parameters may influence characteristics of the created NN, and thus NN performance in detecting ischemic patients, a robustness analysis is necessary before RS&LP can be proposed for clinical use. Thus, aim of this study was to analyze RS&LP robustness to varying values of maximal number of layers, maximal number of initializations and maximal number of confirmations in order to identify the combination of parameters values that provides the best performance in terms of reliability in detecting patients affected by myocardial ischemia.

II. Materials and Methods

A. Ischemia Database

A dataset of serial ECGs from 482 patients, 84 of whom classified as cases (*i.e.,* patients with induced ischemia) and 398 as controls (*i.e.,* patients without ischemia) by expert cardiologists. For each patient, a couple of standard 10-

L. Burattini is with Department of Information Engineering, Università Politecnica delle Marche, via Brecce Bianche 12, 60131, Ancona, Italy (corresponding author; phone: (39) 071 220 4461; fax: (39) 071 220 4224; e-mail: l.burattini@univpm.it).

A. Sbrollini, I. Marcantoni and M. Morettini are with Department of Information Engineering, Università Politecnica delle Marche, via Brecce Bianche 12, 60131, Ancona, Italy (e-mail: a.sbrollini@staff.univpm.it; i.marcantoni@pm.univpm.it; m.morettini@univpm.it).

C.A. Swenne is with Cardiology Department, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands (email: c.a.swenne@lumc.nl).

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

seconds 12-lead ECG was recorded at different time instances. The case patients, whose data belong to the STAFF III database [13], underwent balloon occlusions during elective percutaneous transluminal coronary angioplasty. The first ECG was recorded 3 minutes before the occlusion, in a stable resting condition; the second ECG was recorded 3 minutes after the occlusion, in condition of induced acute myocardial ischemia. The control patients, whose data belong to the digital ECG database of the Leiden University Medical Centre [14], were outpatients of the cardiology department. The first ECG and the second ECG were recorded during routine check-ups, about one year apart, in equal clinical statuses. The present retrospective study was undertaken in compliance with the ethical principles of Helsinki Declaration and approved by the Leiden University Medical Center Medical Ethics Committee.

Serial ECG data of all patients were equally divided into learning dataset and testing dataset. The learning dataset was used to create NNs by using the RS&LP; the testing dataset was used to assess NNs performance in identifying ischemic patients. The learning dataset was further divided into training dataset (80% of the learning dataset) and validation dataset (20% of the learning dataset). Prevalence of cases and controls were maintained in all datasets. Serial ECG data division into learning (training and validation datasets) and testing datasets are reported in Table I.

B. Feature Selection

Each ECG was processed with the Leiden ECG Analysis and Decomposition Software [15] that converts the standard 10-seconds 12-lead ECG into a vectorcardiogram, computes the coherently averaged beat and determines its QRS onset, QRS end and T-wave end. Automatically detected landmarks were visually confirmed and, if needed, manually corrected by two independent ECG analysts. Then, thirteen features were computed [11]: QRS duration (ms); QT interval (ms); QRS maximum amplitude (μV); T-wave maximum amplitude (μV); magnitude of QRS-integral vector (mV·ms); magnitude of T-wave integral vector (mV·ms); QRS complexity (%); T-wave complexity (%); absolute value of QRS-T spatial angle (°) [12], magnitude of the ventricular gradient vector (mV·ms); heart rate (bpm); magnitude of the J-point vector (μV); and T-wave symmetry (%). These features cover the main aspects of cardiac electrical function; thus, if a change in the electrical heart function occurs, one or more features are expected to change accordingly. The thirteen features measured from the firstly acquired ECG were subtracted from the corresponding thirteen features of the secondly acquired ECG. These thirteen serial (*i.e.,* differential) ECG features were used as inputs of the RS&LP.

C. Repeated Structuring & Learning Procedure

RS&LP is a constructive algorithm for automatic creation of supervised fully connected NNs [11]. Number of input and output neurons may vary with application. Briefly, RS&LP implements an iterative procedure including three phases: structuring, learning and confirmation. The procedure starts from an original NN, at the first iteration usually composed by a hidden layer with one neuron. During the structuring phase, the original NN is upgraded in order to get different candidate NNs; upgrading is obtained by adding a neuron in an existing hidden layer or in a new hidden layer. Candidate NNs must respect two structural rules: 1) number of layers cannot exceed the maximal number of layers (NL) a priori set by user; and 2) number of neurons in a layer cannot be larger than number of neurons in the previous layer. During the learning phase, only the newly added neuron (*i.e.,* its weights and bias) is initialized. Initialization is accepted only if, after one only epoch, the new neuron provokes a decrement of the training error. If the performed initialization is not acceptable, the new neuron is re-initialized. Number of initializations of the new neuron cannot exceed the maximal number of initializations (NI) a priori set by the user. All candidate NNs with acceptable initialization of the new neuron are learnt. During the confirmation phase, validation errors of all learnt candidate NNs are compared with the validation error of the original NN. If the validation error of one or more candidate NNs is smaller than the validation error of the original NN, the candidate NN with the smallest validation error becomes the new original NN; if validation errors of all candidate NNs are larger than the validation error of the original NN, the original NN remains confirmed as such. Then, the procedure starts anew. RS&LP stops when at least one of the following stopping criteria occurs: there are no acceptable candidate NNs; the same original NN was confirmed for a number of times not superior to the maximal number of confirmations (NC) a priori set by the user; or the validation error of the original NN is equal to zero. When one of the stopping criteria occurs, the current original NN becomes the final NN, characterized by a specific architecture (ARC; *i.e.,* number of neurons and number of layers, here represented through a vector in which the number of elements represents the number of layers, and the numerical value in each element represents the number of neurons in the corresponding layer).

In the present study, RS&LP was fed with the thirteen serial ECG features and thus counted thirteen neurons in the input layer. In output it provided the probability of a patient being affected by myocardial ischemia, so that a single neuron in the output was considered. While creating NNs on the learning dataset, classes were weighted according with the inverse of their prevalence to compensate for the cases/controls disproportion. Considered neurons had weights and biases ranging between ±1 and a sigmoid activation function. Used optimization algorithm was the scaled-conjugate-gradients algorithm [17] and the validation-based early stopping was used [18] to avoid overfitting.

TABLE I. SERIAL ECG DATA DIVISION INTO LEARNING (TRAINING AND VALIDATION DATASETS) AND TESTING DATASETS.

	Learning Dataset (50%)			**Testing Dataset (50 %)**	**Total**
	Training Dataset (80 %)	**Validation Dataset (20 %)**	**Total**		
Cases	32	10	42	42	84
Controls	161	38	199	199	398
Total	193	48	241	241	482

D. Robustness Study

RS&LP depends on the three parameters a priori set by user, *i.e.,* NL, NI and NC. Since in the initial RS&LP implementation, NL, NI and NC were set at 3, 500 and 10, respectively [11], the final NN created using these parameters values was considered as the reference NN (NN_{REF}) in this study (where for reference with do not intend optimal). RS&LP robustness to varying values these parameters was assessed by changing the value of one parameter at the time. Specifically, to evaluate RS&LP robustness to varying NL, NL was set equal to 1, 2, 4, and 10 while maintaining NI equal to 500 and NC equal to 10; created NN were indicated as NN_{L1}, NN_{L2}, NN_{L4} and NN_{L10}, respectively. To evaluate RS&LP robustness to NI, NI was set equal to 50, 250, 1000 and 1500 while maintaining NL equal to 3 and NC equal to 10; created NN were indicated as NN_{I50}, NN_{I250}, NN_{I1000} and NN_{I1500}, respectively. Eventually, to evaluate RS&LP robustness to varying NC, NC was set equal to 2, 5, 20 and 50 while maintaining NL equal to 3 and NI equal to 500; created NN were indicated as NN_{C2}, NN_{C5}, NN_{C20} and NN_{C50}, respectively. Thus, overall, twelve new final NNs were created and their performance were compared to that of NN_{REF}.

E. Statistics

To avoid initialization dependency, RS&LP was applied 100 times for each combination of parameters values, thus obtaining 100 potential final NNs. For each of them, the area under the curve (AUC) of the receiver operating characteristic (ROC) on the learning dataset was computed the potential final NN with the largest learning AUC became the final NN for that combination of parameters values. Performance of each final NN was quantified by computing the AUC and the 95% confidence intervals (CI) on the testing dataset. ROC curves were compared using the DeLong's tests [19]. Statistical significance (P-value) was set at 0.05.

III. Results

The results of the robustness analysis are reported in Table II. Different final NN architectures were obtained with different combination of parameters values. Maximum number of layers was 4, while number of neurons ranged from 4 to 122. Overall, thirteen NNs were considered; of them, 12 (92%) were characterized by AUC≥80% and 4 (31%) by AUC≥85%. The best performing NNs were NN_{I250} and NN_{C20}, both characterized by AUC=86%.

When varying NL, NN_{L1} was the least complex network (9 neurons in 1 layer) whereas NN_{L4} was the most complex one (22 neurons in 4 layers). When varying NI, N_{I1000} was the least complex network (32 neurons in 3 layers) whereas N_{I250} was the most complex one (70 neurons in 3 layers). When varying NC, N_{C2} was the least complex network (4 neurons in 2 layers) whereas N_{C50} was the most complex one (122 neurons in 3 layers).

Figure 1 shows the testing ROC curves relating to final NNs for all combinations of parameters values. No ROC was found to be statistically different from the one relating to NN_{REF}. When varying NL, all AUCs were higher than 81% (Figure 1.A), with NN_{L4} performing the best (AUC=85%). When varying NI, all AUCs were higher than 80% (Figure 1.B), with NN_{I250} performing the best (AUC= 86%). Eventually, when varying NC, all AUCs were higher than 74% (Figure 1.C), with NN_{C20} performing the best (AUC= 86%).

IV. Discussion

This study analyzed RS&LP robustness to varying values of the three parameters a-priori set by the user, that are NL, NI and NC, in order to identify the combination of these parameters values that provides the best performance when detecting patients affected by myocardial ischemia. Indeed, variations in parameters values lead to different final NNs which may perform differently. Results indicate that RS&LP is very robust. Indeed, despite being characterized by quite different architectures, performances of all NNs were generally similar (similar AUC, similar CI and no statistically different ROC curves; Table II, Figure 1).

RS&LP automatically creates a NN without defining a priori its architecture. Nowadays, the a-priori definition of a NN architecture still requires several trials and computations since no standardized rules have been drawn up. Too simple architectures constituted by few neurons and/or layers may lead to underfitting; on the other hand, too complex architectures constituted by many neurons and/or layers may lead to overfitting. Thus, availability of new constructive procedures able to automatically determine the appropriate NN architecture is desirable. RS&LP determines the NN architecture by uniquely alternating structuring and learning: NN growing is encouraged by trying different NN candidates (avoiding underfitting) and simultaneously discouraged by imposing that the grown NN must perform better (avoiding overfitting). Since the NN can grow following different learning pathways, final NNs with different architectures but performing similarly are possible.

TABLE II. Architecture and performance on the testing dataset of final NNs created with different combinations of NL, NI and NC values.

	NL	NI	NC	ARC	AUC (%)	CI (%)	P
NN_{REF}	3	500	10	[11 9 1]	83	75-91	-
NN_{L1}	1	500	10	[9]	81	72-89	0.73
NN_{L2}	2	500	10	[8 3]	81	73-89	0.75
NN_{L4}	4	500	10	[8 8 3 3]	85	77-92	0.69
NN_{L10}	10	500	10	[13 5]	83	75-91	0.99
NN_{I50}	3	50	10	[16 12 12]	81	73-89	0.76
NN_{I250}	3	250	10	[26 22 22]	86	79-94	0.52
NN_{I1000}	3	1000	10	[12 10 10]	80	72-89	0.71
NN_{I1500}	3	1500	10	[24 12 11]	85	77-93	0.67
NN_{C2}	3	500	2	[3 1]	74	65-83	0.16
NN_{C5}	3	500	5	[8 3 3]	82	74-90	0.86
NN_{C20}	3	500	20	[37 25 25]	86	79-94	0.54
NN_{C50}	3	500	50	[42 40 40]	84	76-92	0.80

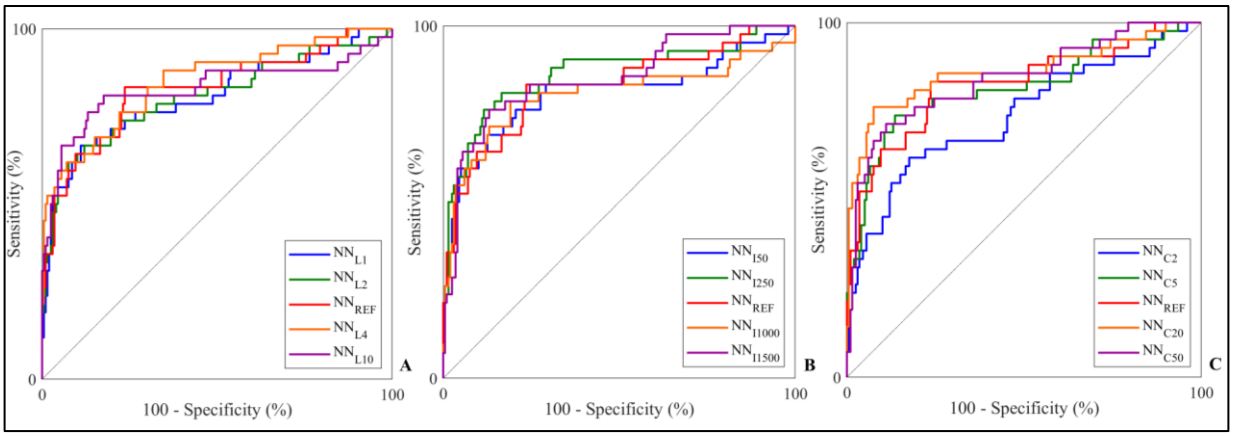

Figure 1. Testing ROC curves relating to final NNs obtained using different combinations of parameters values. Specifically, panels A, B and C depict ROC curves obtained when varying NL, NI and NC, respectively; in all panels, the ROC curve relating to NN_{REF} is depicted in red.

When varying NL, one could expect that all available layers would be filled in the final NN; however, this did not happen for large values of NL (Table II). This finding indicates that, when applying RS&LP, structuring is strongly driven by learning. Once learning obtains good results, RS&LP stops even though more complex NN architectures are potentially reachable, avoiding overfitting. Thus, medium-high values of NL (*i.e.,* NL≥3) are suggested to leave the procedure free to create NN architectures of appropriate complexity. When varying NI, one could expect that the higher the NI value, the higher the number of added neurons, and the more complex the NN architecture; however, this behavior was not observed (Table II). Thus, an intermediate value of NI (around 250) is suggested to limit computational efforts. Finally, when varying NC, NN obtained with medium-high values of NC tended to perform better (Table II), as expected. Thus, an intermediate value of NC (as 10 or 20) is suggested to avoid low performance and high computational efforts.

In conclusion, RS&LP proved to be reliable when applied to detect myocardial ischemia and, thus, can be proposed as a clinical tool, also in emergency situations. Clinical applications require setting of default parameters values that, according to our results, could be NL=3, NI=250 and NC=10.

V. Conclusion

When applied to detect myocardial ischemia, RS&LP proved to be very robust procedure to varying values of NL, NI and NC parameters, which need to be set a priori by the user; suggested values are 3, 250 and 10, respectively.

References

[1] J. Choudhary et al., "Clinical Implications of the ISCHEMIA Trial: Invasive vs Conservative Approach in Stable Coronary Disease," *Curr. Cardiol. Rep.*, vol. 23, no. 5, p. 43, May 2021.

[2] C. C. Haar et al., "An initial exploration of subtraction electrocardiography to detect myocardial ischemia in the prehospital setting," *Ann. Noninvasive Electrocardiol.*, vol. 25, no. 3, May 2020.

[3] K. Ng, S. R. Steinhubl, C. DeFilippi, S. Dey, and W. F. Stewart, "Early detection of heart failure using electronic health records," *Circ. Cardiovasc. Qual. Outcomes*, vol. 9, no. 6, pp. 649–658, Nov. 2016.

[4] W. R. Harlan, A. Graybiel, R. E. Mitchell, A. Oberman, and R. K. Osborne, "Serial electrocardiograms: Their reliability and prognastic validity during a 24-yr period," *J. Chronic Dis.*, vol. 20, no. 11–12, pp. 853–867, Nov. 1967.

[5] B. J. A. Schijvenaars, G. van Herpen, and J. A. Kors, "Intraindividual variability in electrocardiograms," *J. Electrocardiol.*, vol. 41, no. 3, pp. 190–196, May 2008.

[6] L. D. Ostrander, "Serial electrocardiographic findings in a prospective epidemiological study," *Circulation*, vol. 34, no. 6, pp. 1069–1080, Dec. 1966.

[7] D. Marinucci, A. Sbrollini, I. Marcantoni, M. Morettini, C. A. Swenne, and L. Burattini, "Artificial neural network for atrial fibrillation identification in portable devices," *Sensors*, vol. 20, no. 12, p. 3570, Jun. 2020.

[8] X. Wang et al., "Automatic diagnosis of ECG disease based on intelligent simulation modeling," *Biomed. Signal Process. Control.*, vol. 67, p. 102528, May 2021.

[9] F. M. Dias, H. L. M. Monteiro, T. W. Cabral, R. Naji, M. Kuehni, and E. J. da S. Luz, "Arrhythmia classification from single-lead ECG signals using the inter-patient paradigm," *Comput. Methods Programs Biomed.*, vol. 202, p. 105948, Apr. 2021.

[10] P. Li, Y. Hu, and Z.-P. Liu, "Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods," *Biomed. Signal Process. Control.*, vol. 66, p. 102474, Apr. 2021.

[11] A. Sbrollini et al., "Serial electrocardiography to detect newly emerging or aggravating cardiac pathology: a deep-learning approach," *Biomed. Eng. Online*, vol. 18, no. 1, p. 15, Dec. 2019.

[12] A. Sbrollini et al., "Serial ECG analysis: absolute rather than signed changes in the spatial QRS-T angle should be used to detect emerging cardiac pathology," in *Computing in Cardiology*, 2018, pp. 1-4.

[13] S. G. Warren and G. S. Wagner, "The STAFF studies of the first 5minutes of percutaneous coronary angioplasty balloon occlusion in man," *J. Electrocardiol.*, vol. 47, no. 4, pp. 402–407, Jul. 2014.

[14] R. W. Treskes et al., "Performance of ST and ventricular gradient difference vectors in electrocardiographic detection of acute myocardial ischemia," *J. Electrocardiol.*, vol. 48, no. 4, pp. 498–504, Jul. 2015.

[15] H. H. M. Draisma et al., "LEADS: an interactive research oriented ECG/VCG analysis system," in *Computers in Cardiology*, 2005, pp. 515–518.

[16] G. King and L. Zeng, "Logistic Regression in Rare Events Data," *Polit. Anal.*, vol. 9, no. 2, pp. 137–163, Jan. 2001.

[17] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, no. 4, pp. 525–533, Jan. 1993.

[18] L. Prechelt, "Early stopping — but when?," in *Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science*, 2nd ed., G. Montavon, G. B. Orr, and K.-R. Müller, Ed. Berlin: Springer, 2012, pp. 53–67.

[19] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics*, vol. 44, no. 3, p. 837, Sep. 1988.