Feasibility of cardiovascular risk assessment through non-invasive measurements

Feasibility of cardiovascular risk assessment through non-invasive measurements

Pasquale Arpaia[1], Renato Cuocolo[2], Francesco Donnarumma[3], Dario D'Andrea[1], Antonio Esposito[4], Nicola Moccaldi[1], Angela Natalizio[1], and Roberto Prevete[1]

***Abstract*—The present work is a first step in building a wearable system to monitor the heart functionality of a patient and assess the cardiovascular risk by means of non-invasive measurements, such as electrocardiogram (ECG), heart rate, blood oxygenation, and body temperature. Also clinic data obtained by means of a patient interview are taken into account. In this feasibility study, measures from a pre-existing dataset are exploited. They are processed with a machine learning algorithm. Features are first extracted from the measures collected with the wearable sensors. Then, these features are employed together with clinic data to classify the patients health status. A Random Forest classifier was employed and the algorithm was characterized considering different setups. The best accuracy resulted equal to 78.6% in distinguishing three classes of patients, namely healthy, unhealthy non-critical, and unhealthy critical patients.**

I. Introduction

With the increase of elderly people, the need of medical assistance has grown more and more. Continuous medical support in a clinic or hospital is often not feasible in aging populations; yet, collecting medical data is required to maintain a high-quality of life. With the advance of modern machine learning techniques, the analysis of these data could lead to new discoveries and to the possibility of a better patients monitoring and follow-up.

In the framework of a technological development that aims to increase the health and wellness of the population, the Italian project "Indago" has been conceived to develop a wearable system to collect bio-signals from patients directly at their home. It is foreseen a two-fold advantage, (i) avoiding crowded hospitals, which limit the capability of the physicians in patients caring, and (ii) avoiding the psychological stress of the hospital for the patients. Nonetheless, telemonitoring helps in following-up at-risk patients and, when needed, address them to a physician. In order to achieve that, a predictive algorithm for an early diagnosis is foreseen, which analyzes the patient data at the platform input and derives the patient cardiovascular (CV) status. The early diagnosis algorithm aims to assess the CV risk of the monitored patient, relying on bio-signals acquired continuously though a wearable system and clinical data. The latter are obtained through patient interview and clinical examinations. The calculated risk helps in addressing the patient to further examination or it helps the physician in deciding a therapy.

Several contributions are reported in scientific literature with regard to algorithms based on artificial intelligence for the cardiovascular risk assessment. In particular, machine learning techniques are employed to help the physician in formulating a diagnosis or to assess the cardiovascular risk itself. However, these techniques are usually employed to process images, which are results of medical examinations such as a computed axial tomography (CAT). Indeed, fewer contributions are related to the employment of wearable monitoring systems for the cardiovascular state monitoring. In 2010, a smartphone-based platform for the real-time assessment of cardiovascular risk through an electrocardiogram (ECG) has been proposed [1]. The acquired data were also considered for offline elaboration for an Holter-like system. The authors underline the need of a continuous monitoring of at-risk patients because of the mortality related to cardiovascular diseases. Employing an ECG analysis by means of machine learning algorithms, they detect four different kinds of arrhythmia and declare a classification accuracy greater than 90% for three types of arrhythmia, and equal about to the 81% for the fourth type, while the ECG signal is classified as "normal" in 99% of cases. In 2013, a detector of atrial fibrillation has been proposed. It exploits signals related to cardiac vibration, namely a ballistocardiogram (BCG), which is complementary to an ECG analysis [2]. The system is integrated in the patient's bed for the telemonitoring of a cardiovascular disease, and signals are analyzed by means of machine learning. However, the authors themselves underline that the BCG cannot replace an ECG, but it can only add complementary information. The algorithm performance has been reported through the sensibility and specificity, equal to 93.8% and 98.2% respectively. In 2015, researchers from the Turkish University of Inonu have proposed the employment of a "knowledge discovery process" algorithm for stroke prediction. They underlined the importance of this processing for a correct interpretation of available measures, in order to extract synthetic features. The algorithm was based on Artificial Neural Network or Support Vector Machine. The two techniques were compared concluding that the prediction accuracies are 81.82% and 80.38% respectively [3]. The authors foresee the employment of their technique to big datasets in order to enhance the prediction and treatment of stroke, and in their study they considered data from 297 patients. In 2018, a new methodology for the classification of heart health state has been proposed relying on ECG and on evolutionary-genetic algorithms. The "MIH-BIH Arrhythmia database" was taken into account, and the classification discriminates 17 classes of cardiac diseases. The best classifier resulted a Support Vector Machine. The author reports an accuracy equal to 90.20% [4]. In the same year, a wearable system has been proposed for the short-time cardiovascular risk assessment exploiting also the patient's emotional state. A multivariate data analysis is exploited, comprising ECG, heart rate, body temperature,

blood oxygen, and also the "galvanic skin response" (GSR) for the emotional state assessment [5]. The proposed system resulted very similar to the foreseen system architecture within the Indago project. Furthermore, the dataset produced along this study was made publicly available.

In this paper, the feasibility study, based on the classifier Random Forest cited above, and exploiting the available dataset of a machine learning-based algorithm is proposed for the classification of the cardiovascular risk. In particular, Section II discusses the requirements of the Indago system and conceptual design of the algorithm implemented to study the feasibility of the cardiovascular risk assessment procedure, while Section III reports some details about the algorithm implementation and the obtained results. Future steps of the project are finally addressed in the conclusions

II. Method

The proposal of this paper is the implementation of an algorithm for the cardiovascular risk assessment. Data consists of non-invasive measures contained in a pre-existing dataset and in clinic data reported in a related sheet. This aims to demonstrate the feasibility in building a wearable system for the monitoring of cardiovascular diseases. The requirements of the algorithm are thus first reported. They arise from the needs of the project and the foreseen functionalities of the final system under construction. The algorithm for assessing cardiovascular risk receives, as input, data from a continuous monitoring of the patient. The monitoring data are collected from wearable sensors. Instead, another kind of data is also available, namely the above-mentioned clinic data. Clinic data are collected interviewing the patient or recording the results of clinical exams that the patient undertook. Part of these data is collected before the monitoring starts, while another part can be added or updated during the monitoring period.

The data analysis aims to provide a long term risk assessment. The predicting capability of the algorithm involves the possibility to foresee a cardiovascular risk from the patient data, acquired through the telemonitoring system. If an increase of the cardiovascular risk is foreseen, the patient is addressed to a physician for further analysis and/or for the prescription of a therapy. Nonetheless, if the risk remains unchanged or it decreases, the patient is suggested to undertake classical routine exams. The main output of the system, namely of the predictive algorithm, is an estimation of patient long term cardiovascular risk. In accordance to the most recent European guidelines [6], this value should be expressed as a categorical class: low-risk, moderate-risk, high-risk and very-high risk. Fig. 1 reports some details about these four risk categories according to the mentioned guidelines.

Thanks to the detection of eventual anomalies in the monitored data, patient current status can also be reported. Moreover, when the algorithm predicts an increase of cardiovascular risk, the physician may receive an alert. Finally, in case of an unforeseen cardiovascular event, the recorded bio-signals can help in obtaining a prompt diagnosis. All this information, in turn, could help improve the clinical management of the

Very high-risk	Subjects with any of the following: • Documented CVD, clinical or unequivocal on imaging. Documented clinical CVD includes previous AMI, ACS, coronary revascularization and other arterial revascularization procedures, stroke and TIA, aortic aneurysm and PAD. Unequivocally documented CVD on imaging includes significant plaque on coronary angiography or carotid ultrasound. It does NOT include some increase in continuous imaging parameters such as intima–media thickness of the carotid artery. • DM with target organ damage such as proteinuria or with a major risk factor such as smoking or marked hypercholesterolaemia or marked hypertension. • Severe CKD (GFR <30 mL/min/1.73 m2). • A calculated SCORE ≥10%.
High-risk	Subjects with: • Markedly elevated single risk factors, in particular cholesterol >8 mmol/L (>310 mg/dL) (e.g. in familial hypercholesterolaemia) or BP ≥180/110 mmHg. • Most other people with DM (with the exception of young people with type 1 DM and without major risk factors that may be at low or moderate risk). • Moderate CKD (GFR 30–59 mL/min/1.73 m²). • A calculated SCORE ≥5% and <10%.
Moderate-risk	SCORE is ≥1% and <5% at 10 years. Many middle-aged subjects belong to this category.
Low-risk	SCORE <1%.

Figure 1: Risk categories. ACS: acute coronary syndrome; AMI: acute myocardial infarction; BP: blood pressure; CKD: chronic kidney disease; DM: diabetes mellitus; GFR: glomerular filtration rate; PAD: peripheral artery disease; SCORE: systematic coronary risk estimation; TIA: transient ischemic attack

subjects. Details about the algorithm implemented in this first study are reported in the following subsection.

A. Algorithm implementation

Two main steps can be identified in the algorithm: features extraction and classification. For features extraction, data from wearable sensors is taken into account, together with data resulting from a patient interview (clinic data). Non-invasive measures from the wearable sensors are processed to obtain the features as better described below, while clinic data is employed as is. Then, subjects data are separated to train the classifier and then validate it. The algorithm architecture is shown in Fig. 2.

Several features can be considered for each patient, also depending of the availability of a type of feature for all patients. Considering data from wearable sensors first, ECG signals are processed with the Pan-Tompkins algorithm [7] to obtain the distances between consecutive R-peaks. This information is exploited to calculate 12 features: heart rate, maximum, minimum and average RR distance, standard devi-

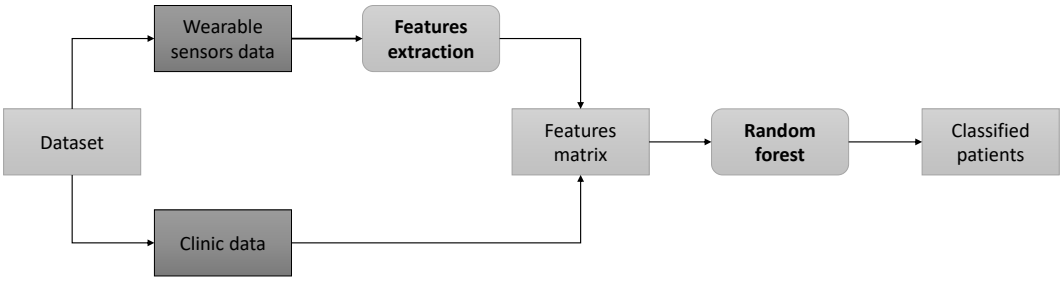

Figure 2: Architecture of the proposed algorithm for the cardiovascular risk assessment. Part of data is processed to extract features, while the remaining part of data is considered as features themselves. Then the Random Forest classifier is trained and tested properly splitting the data.

ation of RR distances (SDNN), the square root of the mean squared differences of consecutive RR (rMSSD), consecutive NN intervals differences which is higher than 50 ms (NN50), NN50 by the total number of NN intervals (pNN50), HRV triangular index (HRVI), the imaginary ellipse axis SD1 and SD2 of the Poincaré plot, and their ratio SD1/SD2. These features can be calculated from an ECG signal during a resting phase, during a 6-minute walking test, and when the patient was watching a video (2 videos maximum are available). Other features arise from the remaining measures, available for each kind of trial, namely body temperature, blood oxygenation (SpO2), and heart rate. The synthetic features taken into account are the mean body temperature, and the mean and standard deviation of SpO2 and heart rate, yielding a total of 10 more features. Finally, 24 features are the clinic data listed here for clarity: age, number of stents, presence of bypass, smoking status, metabolic syndrome status, diabetes status, angina or heart attack in a 1st degree relative, chronic kidney disease (stage 3. 4 or 5), atrial fibrillation, on blood pressure treatment, migraines, rheumatoid arthritis, SLE, illness, atypical anti-psychotic medication, regular steroid tablets, erectile disfunction, height, weight, 3 measures of systolic blood pressure, diastolic pressure. The features matrix is thus built considering the features on columns and the patients on rows. This is the input of the classifier. The matrix was obtained with Matlab by processing data from wearable sensors and concatenating clinic data. The output file is a '.csv' file.

The classifier that was implemented is a multi-class Random Forest, which is a classical machine learning technique [8]. The Random Forest has been considered as the optimal classifier, as suggested by the literature, among different types of classifiers that were studied, such as decision forest, logistic regression, decision jungle, and neural networks.This classifier was developed in Python, and it receives as input the previously generated '.csv' file. The Python script simply loads the features matrix and then it adopts as classification methods, relying on the Random Forest technique the hold-out method. It consists in splitting randomly the dataset in 50% for training the model and 50% for testing it and obtaining a classification accuracy value. The array of label was uploaded too, and it was exploited for both training and testing of the Random Forest classifier. The results of the methods application are reported in the following section.

III. Experimental results

The algorithm discussed in the previous section was implemented and the dataset received from the authors of [5] was employed for testing. The dataset is described in the following subsection, while the results and their discussion are reported in a further subsection considering different training and testing data choices.

A. Dataset description

The employed dataset contains data from 30 patients. For each patient, the clinic data listed before is reported in a resume sheet, while recorded data from wearable sensors is furnished in separated files. Each file contains the ECG signal, GSR, pulse, body temperature, blood oxygenation (SpO2), and a time-stamp. The signals were sampled at 290 Sa/s during acquisition. The Galvanic Skin Response (GSR) has been ignored in this first implementation, as well as the time-stamp, which resulted useless for our needs. The different files are related to 4 different states during acquisition: 10 minutes resting state, 6-minute walking test (6MWT), watching two 6 minutes videos. The videos were mostly meant for the emotional response. However, our study also took into account the data from wearable sensors acquired when watching a video. The 6MWT was then discarded because all unhealthy patients refused to undertook it. The effectively considered patients depend on some choices related to the need to balance the different classes of patients for training and test, hence this is better discussed in the results subsection.

The label associated to each patient distinguishes healthy patients from unhealthy ones, and then again critical and non-critical unhealthy patients. Health status was assessed by a cardiac physicians after some examination. These labels are employed for both classification and evaluation phases of the classifier. The dataset taken into account contains 16 healthy patients, 10 unhealthy non-critical patients, and only 4 unhealthy critical patients.

B. Results and discussion

The execution of the algorithm yields different accuracy values according to the choice of the data for training and testing. The different setups are hence reported along with the achieved accuracies. First, aiming to balance the dataset, the considered patients were 8 healthy, 8 unhealthy non-critical and 4 unhealthy critical patients. Unfortunately, the number of unhealthy critical patients is strongly limited. However, considering less patients for the remaining classes would have meant to limit the number of data for training and testing, which is already small. Hence, it was not possible to consider a perfectly balanced sub-dataset. The classification conducted in this first setup yields an accuracy equal to 70% when the ECG, body temperature, blood oxygenation (SpO2) and pulse are considered in resting state. The accuracy is still 70% if clinic data is added in the features matrix. Instead, if only clinic data is employed in this setup, the resulting accuracy drops to 60%. Hence, this indicates that clinic data adds no relevant information to the classification of patients in this configuration.

The second choice regarded the possibility to employ as much data as possible while considering also measurements during video watching. Hence, 28 patients were considered, while 2 patients had to be discarded because data during video watching was not present for them. In this case, the considered patients were 16 healthy patients, 9 unhealthy non-critical patients, and 3 unhealthy critical ones. The accuracy resulted equal to 78.6% in two conditions, with only clinic data, or with clinic data and all measures during resting state and video watching. Instead, if only the resting state is considered, and no clinic data is employed, the accuracy drops to 71.4%. The same value is obtained when only data acquired during video watching is considered.

Finally, all patients were considered, accepting to discard data acquired during video watching. These data were not present for the 2 patients that were previously discarded. In this setup, considering only the resting state, the yielded accuracy is 66.7%. This value remains unchanged if clinic data are added as features. Instead, if only clinic data are taken into account, the accuracy results equal to 73.3%.

All results are resumed in Tab. I. They depict a situation of clinic data leading to the best accuracy value, when the classes of patients taken into account are not balanced, and in particular there are mostly healthy patients. Instead, when aiming to balance the classes, the best result is achieved thanks to the data from wearable sensors acquired during resting state. Finally, the results seem to suggest that resting state and video watching are equivalent conditions. This is reasonable because the galvanic skin response (GSR) was not considered in our study. Instead, the GSR could give further information about the emotional state of the patient, which would eventually lead to a classification improvement, as the author of [5] report. This investigation is addressed to a next work.

Setup	Data	Accuracy
1st	resting state	70%
1st	resting state, clinic data	70%
1st	clinic data	60%
2nd	resting state	71.4%
2nd	video watching	71.4%
2nd	clinic data	78.6%
2nd	resting state, video watching, clinic data	78.6%
3rd	resting state	66.7%
3rd	resting state, clinic data	66.7%
3rd	clinic data	73.3%

Table I: Classification accuracy results for the different setups discussed within the manuscript and for different data considered as features.

IV. Conclusions

In this paper, an algorithm based on Random Forest classification for the cardiovascular risk assessment was proposed. The input consists of data recorded from sensors that the patient can wear and on clinic data obtained interviewing the patient. After features extraction, the patient are classified through a Random Forest algorithm. The best result in terms of accuracy is obtained exploiting an unbalanced dataset with mostly healthy patients, yielding a 78.6% accuracy in distinguishing 3 different classes. This is explained with the fact that the number of patients is too small to properly train and test the classifier, hence the best results is obtained when the whole dataset is employed. However, this also means that over-fitting problems could be present, and the classifier could poorly classify incoming data from new patients.

This feasibility study is a first step in building a smart wearable system for the monitoring of the cardiovascular condition of patients.

References

[1] J. J. Oresko, Z. Jin, J. Cheng, S. Huang, Y. Sun, H. Duschl, and A. C. Cheng, "A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 734–740, 2010.

[2] C. Bruser, J. Diesel, M. D. Zink, S. Winter, P. Schauerte, and S. Leonhardt, "Automatic detection of atrial fibrillation in cardiac vibration signals," *IEEE journal of biomedical and health informatics*, vol. 17, no. 1, pp. 162–171, 2013.

[3] C. Colak, E. Karaman, and M. G. Turtay, "Application of knowledge discovery process on the prediction of stroke," *Computer methods and programs in biomedicine*, vol. 119, no. 3, pp. 181–185, 2015.

[4] P. Pławiak, "Novel methodology of cardiac health recognition based on ecg signals and evolutionary-neural system," *Expert Systems with Applications*, vol. 92, pp. 334–349, 2018.

[5] F. P. Akbulut and A. Akan, "A smart wearable system for short-term cardiovascular risk assessment with emotional dynamics," *Measurement*, vol. 128, pp. 237–246, 2018.

[6] M. F. Piepoli, A. W. Hoes, S. Agewall, C. Albus, C. Brotons, A. L. Catapano, M.-T. Cooney, U. Corra, B. Cosyns, C. Deaton, *et al.*, "2016 european guidelines on cardiovascular disease prevention in clinical practice: The sixth joint task force of the european society of cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of 10 societies and by invited experts)

developed with the special contribution of the european association for cardiovascular prevention & rehabilitation (eacpr)," *European heart journal*, vol. 37, no. 29, pp. 2315–2381, 2016.

[7] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE Trans. Biomed. Eng*, vol. 32, no. 3, pp. 230–236, 1985.

[8] L. Rokach, "Decision forest: Twenty years of research," *Information Fusion*, vol. 27, pp. 111–125, 2016.