Coronary Health Index (CHI) as a Determinant for Arterial Stenosis, derived Using PPG and ECG Signals

Poulomi Pal, Manjunatha Mahadevappa

School of Medical Science and Technology, Indian Institute of Technology Kharagpur, India

Abstract

Cardiovascular disease (CVD) patients were targeted in this study to assess the performance of the proposed methodology in segregating patients having stenosis and also identifying the principal diseased coronary artery using PPG and ECG signals. These signals were acquired simultaneously from patients in cardiology department . After pre-processing the signals, dicrotic notch region of PPG and S-T segment of ECG, within each cardiac cycle was extracted as a template. A new fused segment was generated from two templates by a proposed algorithm. Utilizing statistics on three templates we defined the term Coronary Health Index (CHI) to evaluate the status of coronary arteries. Setting CHI thresholding values, healthy and stenosed artery were differentiated. Using CHI values from patients with stenosis, the classification of arteries (LAD, RCA, and LCx) was performed using Graph Attentive Convolution Network (GACN). Among 408 CVD patients 256 had occlusion in either LAD or RCA or LCx. Binary classification among presence and absence of stenosis was carried out with 0.92 accuracy, 0.91 recall, 0.91 precision, 0.90 specificity, and 0.92 F-score. Identification of stenosed artery was measured with Kappa score (0.89) and Youden's J statistic value (0.84). AUC(0.93) and AP(0.92) values from ROC and PRC curves, respectively. The CHI derived from the PPG and ECG signals could be able to study stenosis in non-invasive, easy and cost-effective manner.

1. Introduction

Coronary artery disease (CAD) is a prevalent cardiovascular disease affecting people worldwide. CAD is a medical condition where a plaque (generally constitutes of fat, calcium, cholesterol, etc.) formation takes place inside the coronary arteries of heart. This prevents supply of blood (oxygenated) to myocardium, thus affecting essential functioning of the heart. It is also known as IHD or ischemic heart disease. As per health statistics, there was an increase in deaths due to CAD or IHD by 74.9% in 2017 when compared to 1990 [1]. According to global predictions this disease will affect 1845 people in 1 lakh in 2030 from 1655 people in present [2]. The cardiac patients admitted to hospital were the cohorts in acquiring ECG and PPG signals. Three major coronary arteries left right coronary artery (RCA), anterior descending coronary artery (LAD), and left circumflex coronary artery (LCx) are one of the most important arteries which get affected most due to stenosis as per the medical literature and cardiologists suggestion. So, these arteries were focused in this study to find the possibility of coronary artery identification in CAD patients.

Biomedical researches are being carried out to detect CAD using physiological signals. ECG signals have been used for quite a long time in medical field for all types of patients. Photoplethysmography signals are gaining attention in predicting status of cardiological functions. There are works where ECG and PPG signals were individually used to detect CAD.

Yao et al. used residual network to build an ML algorithm to detect CAD from ECG signal features with 0.96 accuracy, 0.95 sensitivity, and 0.96 specificity [3]. Ventricular polarization measurements in QRS, S and T waves of ECG signals were used to classify CAD group with sensitivity and specificity 0.63 and 0.75, respectively. Li et al used dual inputs ECG and PCG signals to extract features and use them in a deep learning model to identify CAD patients with specificity of 0.89 [4]. Dey et al. used SVM and Wavelet transform to differentiate CAD patients with an accuracy of 0.89 on PPG signal online data repository [5]. The time domain analysis of PPG signals to identify CAD patients using SVM was done by Paradkar et al. with the 0.85 sensitivity of and 0.78 specificity [6]. Ouyang et al. distinguished CAD and pre-CAD from healthy subjects using PPG signals from ear lobe, finger and toe tip with an accuracy of 0.83 [7].

The aim was to acquire PPG and ECG signals were simultaneously from CVD patients visiting hospital and distinguish CAD and non-CAD patients. ECG and PPG templates were extracted from these processed signals. Fused template was derived from these two templates. After finding the coronary health index (CHI) values we segregated the patients using the graph attentive convolution neural netwrok (GACNN). Later among the CAD patients we classified the patients on the basis of the coronary artery location weather it is in LAD, RCA and LCx. We analysed the CAD patients using this proposed CHI values.

2. Methodology

2.1. Data

The two signals (ECG and PPG) were acquired at the same time from CVD patients visiting hospital for treatment. The coronary artery angiogram (CAG) reports was considered as the *gold standard data* for diagnosing CAD patients. Ethical clearance from Institutional Ethical Committee (IEC) was obtained from the Indian Institute of Technology Kharagpur, with vide reference no. : IIT/ SRIC/DR/2019, dated 06-11-2019 and Medical College and Hospital, Kolkata with vide reference no. : MC/Kol/IEC/Non-spon/139/09-2018, dated 10-11-2018. The inclusion and exclusion of cohorts were followed as decided by IEC. Before data collection, the participants were asked to give attestation after going through the IEC consent forms as per ethical norms.

PPG signals were collected using the experimental setup consisting sensor - IR Plethysmograph Velcro Strap (MLT 1020 PPG, AD Instruments, Sydney, Australia) and three clamp electrodes of red, green, and black colors for collecting PPG and ECG signals, respectively. A bio-amplifier (Dual Bio-AMP-FE 232 AD Instruments, Sydney, Australia) is used to increase signals' amplitude. In the end, a DAQ (Power Lab r 8/35, ML135, AD Instruments, Sydney, Australia) recorded the signals at resolution of 16-bit, zero phase, and 2 KHz acquisition frequency. A high pass filter with cut off 0.2 Hz was used to prevent the undesired low frequencies. Notch filter of 50 Hz removed power line interferences. High- frequency noise was removed by 50 Hz Butter-worth filter of order=8 with 1000Hz sampling frequency.

The demographic details of recruited patients were age, height, weight, systolic pressure, diastolic pressure, pulse rate of the cohorts were 56.11 ± 12.29 yrs, 166.34 ± 6.62 cm, 61.75 ± 13.65 Kg, 121.06 ± 18.70 mm of Hg, 78.03 ± 11.47 mm of Hg, 77.48 ± 11.38 bpm, respectively.

2.2. Template extraction

Pre-processed PPG and ECG signals were used for each patient to derive the coronary health index (CHI). According to literature, ST segment of ECG and dicrotic notch of PPG were referred for CAD detection in cardiac patients. So, both the signals were segmented equally, observing the cardiac cycles. For ECG, the ST segment was extracted locating the T wave and J point in the segment. This part of the ECG signal was stored as the ECG template. In case of PPG segment, the catacrotic phase was targeted. The area including diastolic peak and dicrotic notch was extracted and stored in PPG template. These templates were used to generate the third template - fused template. An algorithm was designed to generate this fused segment. ECG and PPG templates were used as the inputs. Spline interpolation was used to find the third template. Single 1D matrix was obtained using concatenation operation on interpolated template and two input templates. Now for each patient three templates were obtained from PPG and ECG signals.

The functions involved to find CHI for each patient included maximum likelihood estimation (MLE), correlation (r) and covariance (c). MLE was found from the fused templates to calculate the conditional probability in their probability distributions. r and c was found among the PPG and ECG templates to calculate the commonality and differences, respectively. The MLE ($\hat{\theta}$) of a parameter $\theta_\circ$ was found using (1)

$$\hat{\theta} = arg(\underset{\theta \in \Theta}{max}) ln \left[L\left(\theta; \xi\right) \right] \tag{1}$$

where, Θ is the parameter set, ξ is sample, $L(\theta; \xi)$ is likelihood of sample, $arg\ max$ gives maximum value of log likelihood estimation of the sample. The correlation (r) of PPG templates x and ECG templates y was found using (2)

$$r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x^2)][n\Sigma y^2 - (\Sigma y^2)]}} \tag{2}$$

and the correlation (c) of PPG template X and ECG template Y was found using (3)

$$c = \frac{\Sigma(X_i - \bar{X})(Y_j - \bar{Y})}{n} \tag{3}$$

where, n is the number of templates in a cardiac cycle. The values of MLE, r, and c were implemented in (4) to get the CHI value of one patient.

$$CHI = \frac{MLE \times r}{c} \tag{4}$$

These CHI values were used to distinguish CAD patients from non-CAD patients. It was seen the healthy subjects produced CHI above 60 and the CAD patients had CHI value below 60. After identifying CAD patients we used their CHI values to find the coronary artery which has maximum stenosis.

2.3. Classification Network

Using CHI values from patients with stenosis. The classification among LAD, RCA, and LCx was performed using Graph Attentive Convolution Network (GACN). In

GACN the node values were the CHI values. The clinical information among the patients were utilized to derive the edge relationships.

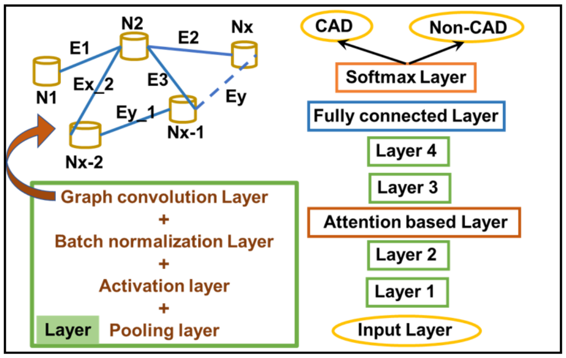

Figure 1. The network topology

The network comprised 4 layers of graph convolution, attention based layer, fully connected layer, softmax layer, and classification layer. Each layer had the graph convolution layer, batch normalization, activation layer, and pooling layer. Self-attention principle was used in the attention based layer. Adam optimization was used as the training algorithm. The mini-batch size was 15, learning rate was 0.001, and momentum was initialized as 0.9. All the coding was carried out in Python platform. The hardware setup was ASUS TUF FX504 system, with GPU of NVIDIA Geforce GTX.

3. Result

The presence of arterial stenosis was confirmed using the proposed GACNN. The input data was divided into training, testing, and validation, data sets in the ratio 70:15:15. The CHI values were observed but performance evaluation was conducted using confusion matrix. 10-fold cross-validation was used for verification of statistical significance. This binary classification between CAD and non-CAD was carried out with 0.92 accuracy, 0.91 recall, 0.91 precision, 0.90 specificity, and 0.92 F-score. Initially for the robustnes of the network ablation studies were conducted using different number of layers, various activation functions, and types of pooling layers. The final network for binary segregation was selected on the basis of best results.

The comparison of the performance of GACNN with other established networks was also done. Here, AlexNet, GoogleNet, InceptionNet, ResNet, and VGG-Net were simultaneously evaluated on the same input data for deriving accuracy, recall and F-score. The hyper-parameter specifications and the hardware platform was kept same throughout. From Table 1 it been be seen that GACNN had produced highest result in comparison to the other networks.

The graph containing mirror Receiver Operating Characteristics (ROC) curves and Precision Recall curves (PR) were studied to derive the Area under the ROC curve (AUC) and Area under the PRC curve (AP) values.

Table 1. Comparative studies with other networks.

Networks	Accuracy	Recall	F-score
GoogleNet	0.87	0.86	0.85
InceptionNet	0.84	0.85	0.84
ResNet	0.88	0.87	0.86
VGG-16Net	0.89	0.86	0.87
GACNN	0.92	0.91	0.90

After analysis of the graph highest values of AUC(0.93) and AP(0.92) values from ROC and PRC curves, were found for GACNN, respectively (from Fig 2).

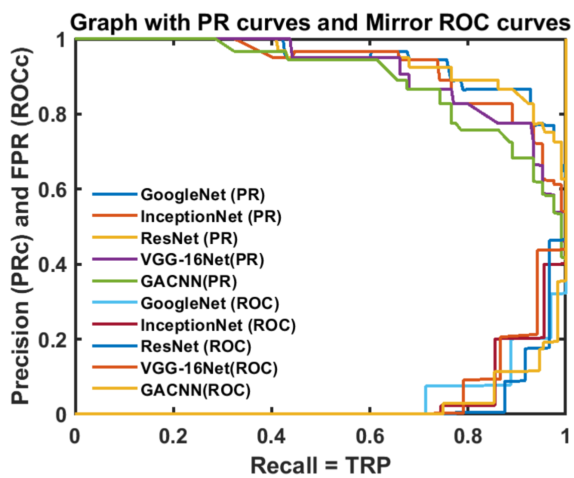

Figure 2. Combined PR curve (PRc) and mirror ROC curve (ROCc) plot for comparing performance of the proposed with standard existing networks

The state-of-the art studies was done with respect to differentiating CAD and non-CAD patients, in terms of the techniques utilised and resulting accuracy, as illustrated in Table 2. Among the CAD patients further classification was carried out concentrating the stenosis in one of the coronary artery. Majorly three important coronary arteries was studied - LAD, RCA, and LCx. The same GACNN architecture was implemented but with less number of inputs. The training, validation, and testing data sets were in 70:20:10 ratio. Cross-validation was conducted in 5-folds. Here, the results were found from the confusion matrix in terms of Cohen Kappa score (0.89), Mico-Avg F1-score (0.88), and Youden's J statistic value (0.84).

4. Discussion

The demonstration of the effectiveness in finding CHI values and using them as inputs to GACNN was prominent from the results. It was the first time where such kind of index has been introduced. Quantitative comparison was made with past works from literature survey.

Table 2. State-of-the-art studies.

Researchers	Techniques	Accuracy
Yao *et al.* [3]	ML	0.96
Paradhkar *et al.* [4]	SVM	0.85
Ouyang *et al.* [5]	signal processing	0.83
Li *et al.* [6]	DL	0.89
Lin *et al.* [7]	signal processing	0.65
Proposed work	GACNN	0.92

The various techniques implemented on PPG signals and the result obtained has been displayed in Table 2. Signal procesing, machine learning (ML), and deep learning were used in [3] to [7]. The present work also involved DL, in network GACNN but the inputs were CHI values derived from statistical analysis of signals. This procedure helped in achieving an accuracy of 0.92. Hence, proving its superiority over the previous works.

During data collection the CAD patients with other comorbidities were excluded due as per ethical norms. Moreover the CAD patients with stenosis in any one of the artery (LAD, RCA, and LCx) were targeted more. Those patients having stenosis in more than one artery were not considered for the study. Thus these were the limitations to the study.

5. Conclusion

The PPG signals were collected from cardiovascular patients admitted in the hospital for treatment. After preprocessing the signals individually templates were extracted from each of the cardiac cycle of both the signals. Using the designed algorithm a fused template was evolved. Statistical tools were implemented to find the CHI using the three templates. Based on these CHI values classification of CAD patients were done using the proposed GACNN. The results were produced with 0.92 accuracy, 0.91 recall, 0.91 precision, 0.90 specificity, and 0.92 F-score. Among the CAD patients, classification of the patients based on the stenosed artery among LAD, RCA, and LCx was executed using the same CHI values and GACNN. Thus, it was observed that CHI values were indicators of CAD non-invasively requiring less time.

Acknowledgments

Authors acknowledge the Indian Institute of Technology Kharagpur for helping and supporting to execute this study. They also gratefully acknowlegde Dr. Bhabani Prasad Chattopadhyay, an Associate Professor, cardiologists, and medical staff working in the Department of Cardiology at Medical College and Hospital, Kolkata. A noteworthy contribution was done by the CVD patients who willingly gave the ECG and PPG signal data. [2–9]

References

[1] Nowbar AN, Gitto M, Howard JP, Francis DP, Al-Lamee R. Mortality from ischemic heart disease: Analysis of data from the world health organization and coronary artery disease risk factors from ncd risk factor collaboration. Circulation cardiovascular quality and outcomes 2019;12(6):e005375.

[2] Wilhelms M, Dossel O, Seemann G. In silico investigation of electrically silent acute cardiac ischemia in the human ventricles. IEEE transactions on biomedical engineering 2011; 58(10):2961–2964.

[3] Yao L, Liu C, Li P, Wang J, Liu Y, Li W, Wang X, Li H, Zhang H. Enhanced automated diagnosis of coronary artery disease using features extracted from qt interval time series and st–t waveform. IEEE Access 2020;8:129510–129524.

[4] Paradkar N, Roy Chowdhury S. Coronary artery disease detection using photoplethysmography. In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2017; 100–103.

[5] Ouyang V, Ma B, Pignatelli N, Sengupta S, Sengupta P, Mungulmare K, Fletcher RR. The use of multi-site photoplethysmography (PPG) as a screening tool for coronary arterial disease and atherosclerosis. Physiological Measurement jun 2021;42(6):064006.

[6] Li H, Wang X, Liu C, Wang Y, Li P, Tang H, Yao L, Zhang H. Dual-input neural network integrating feature extraction and deep learning for coronary artery disease detection using electrocardiogram and phonocardiogram. IEEE Access 2019;7:146457–146469.

[7] Lin Lj, Tang Fk, Hua N, Lu H. Contrast between magnetocardiography and electrocardiography for the early diagnosis of coronary artery disease in patients with acute chest pain. In 2011 Computing in Cardiology. IEEE, 2011; 641–643.

[8] Dey S, Biswas S, Pal A, Mukherjee A, Garain U, Mandana K. Cad patient classification using mimic-ii. In eHealth 360°. Springer, 2017; 370–375.

[9] Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G. The graph neural network model. IEEE Transactions on Neural Networks 2009;20(1):61–80.

Address for correspondence:

Manjunatha Mahadevappa
School of Medical Science and Technology, Indian Institute of Technology Kharagpur, Kharagpur-721302, West Bengal, India.
mmaha2@smst.iitkgp.ac.in