Computers in Biology and Medicine 111 (2019) 103346

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Machine learning-based coronary artery disease diagnosis: A comprehensive review

Roohallah Alizadehsani[a,*], Moloud Abdar[b], Mohamad Roshanzamir[c], Abbas Khosravi[a], Parham M. Kebria[a], Fahime Khozeimeh[d], Saeid Nahavandi[a], Nizal Sarrafzadegan[e,i], U. Rajendra Acharya[f,g,h]

[a] Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Australia
[b] Département d'informatique, Université du Québec à Montréal, Montréal, Québec, Canada
[c] Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran
[d] Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
[e] Faculty of Medicine, SPPH, University of British Columbia, Vancouver, BC, Canada
[f] Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore
[g] Department of Biomedical Engineering, School of Science and Technology, Singapore University of Social Sciences, Singapore
[h] Department of Biomedical Engineering, Faculty of Engineering, University of Malaya, Malaysia
[i] Isfahan Cardiovascular Research Center, Cardiovascular Research Institute, Isfahan University of Medical Sciences, Khorram Ave, Isfahan, Iran

ARTICLE INFO

Keywords:
CAD diagnosis
Machine learning
Data mining
Feature selection

ABSTRACT

Coronary artery disease (CAD) is the most common cardiovascular disease (CVD) and often leads to a heart attack. It annually causes millions of deaths and billions of dollars in financial losses worldwide. Angiography, which is invasive and risky, is the standard procedure for diagnosing CAD. Alternatively, machine learning (ML) techniques have been widely used in the literature as fast, affordable, and noninvasive approaches for CAD detection. The results that have been published on ML-based CAD diagnosis differ substantially in terms of the analyzed datasets, sample sizes, features, location of data collection, performance metrics, and applied ML techniques. Due to these fundamental differences, achievements in the literature cannot be generalized. This paper conducts a comprehensive and multifaceted review of all relevant studies that were published between 1992 and 2019 for ML-based CAD diagnosis. The impacts of various factors, such as dataset characteristics (geographical location, sample size, features, and the stenosis of each coronary artery) and applied ML techniques (feature selection, performance metrics, and method) are investigated in detail. Finally, the important challenges and shortcomings of ML-based CAD diagnosis are discussed.

1. Introduction

In the past few decades, artificial intelligence (AI) has been widely used for various tasks [1], such as driving cars [2], various computer games [3,4], and healthcare [5–8]. In AI, these tasks are commonly done by machine learning (ML) algorithms. AI includes a variety of ML algorithms. ML techniques can be classified into categories; additional information about this matter is provided in the supplementary information (see the “Main types of ML/DM techniques” part of the Supplementary Information) [9–18]. There are many ML methods including support vector machines (SVMs) [19], artificial neural networks (ANNs) [20], decision trees (DTs) [21], Naïve Bayes [22], k-nearest neighbors (KNN) [23], and K-means [24]. Each of these techniques have their strengths and weaknesses, which are listed in Supplementary Table S1. These methods have been widely investigated in broad areas such as medical applications (e.g., liver disease [25,26], electrocardiogram (ECG) signals of the human heart [27,28], Parkinson's disease [29,30], and skin disease [31–33]) for purposes such as screening, risk stratification, prediction, and assisted decision-making [34–37]. The number of research differ due to the mortality rates among various diseases. In other words, more important diseases have attracted more research and attention.

The World Health Organization (WHO) has listed cardiovascular diseases (CVDs) as the leading cause of death around the globe [38] (see Fig. 1b). As the most common type of CVD, coronary artery disease (CAD) occurs when there is an obstruction (of more than 50%) in at least one of the coronary arteries [39]. There are *three* major arteries of the heart: (i) left anterior descending artery (LAD), (ii) left circumflex

* Corresponding author.
E-mail address: ralizadehsani@deakin.edu.au (R. Alizadehsani).

https://doi.org/10.1016/j.compbiomed.2019.103346
Received 7 May 2019; Received in revised form 26 June 2019; Accepted 26 June 2019

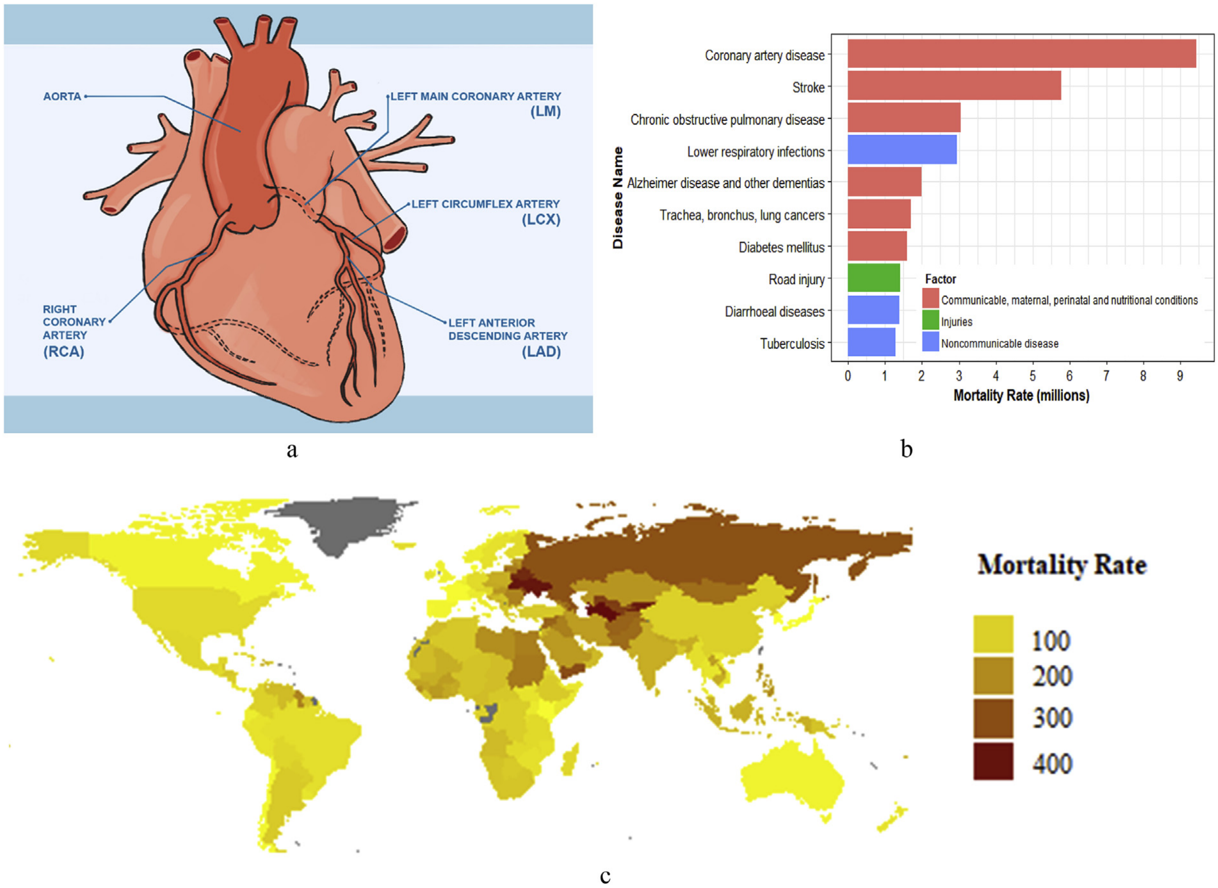


Fig. 1. Importance of investigating CAD. a) The three main arteries of heart are shown in this figure. They are LAD, LCX, and RCA. The left main coronary artery arises from the aorta. It typically runs for 10–25 mm and subsequently divides into two branches: LAD and LCX. If the stenosis in at least one of these arteries exceeds 50%, CAD occurs [39]. Since the left main artery is larger than LAD and LCX, the probability of its stenosis is much lower compared to LAD and LCX [39]. **b)** The top cause of death in the world is coronary artery disease [40]. Its mortality exceeds the second highest cause of death (stroke) by a factor of about two, and the corresponding number of deaths is nearly equal to the sum of the deaths due to the six lower causes in this diagram [39]. **c)** This map shows the number of CAD deaths per 100,000 in various countries, as reported by the WHO [41]. The highest death rates occur in the independent states of the former Soviet Union, the Middle East, and North Africa [39].

artery (LCX) and (iii) right coronary artery (RCA) (see Fig. 1a). CAD mortality rates differ among the regions of the world, as shown in Fig. 1c. These rates are much higher in Asia, Russia, and the Middle East compared to the rest of the world [39]. Hence, providing precise diagnostic and preventive methods can have a momentous impact on reducing mortality due to CAD.

Early detection of CAD is critical to avoid further increase in the risk. Coronary angiography is required to conclusively diagnose CAD [39]. However, it is invasive and may lead to various complications, such as artery dissection, arrhythmia and even death. Moreover, image-based detection techniques are costly and not applicable for screening large populations, especially in developing countries. Due to these shortcomings and the life-threatening nature of angiography, researchers have been continuously looking for noninvasive, economical, fast, and reliable techniques for early detection of CAD. ML algorithms are some of the techniques used for this purpose.

2. ML for CAD diagnosis

ML-based techniques have been successfully applied on various types of CAD datasets [42–48]. These algorithms have demonstrated promising performance in the detection and treatment of CAD. This study comprehensively reviews how, when, and where ML techniques have been applied for CAD detection. By examining the works that have been published since 1992, this study aims at addressing the following: (a) various ML techniques that have been applied and their pros and cons, (b) the dataset characteristics (features and size) and how they impact the overall performance, and (c) performance achieved by each ML method for each dataset. The findings obtained in this study may assist researchers, physicians, governments and patients to take necessary steps to improve the quality of life. Both researchers and physicians are seeking economical, accurate and fast CAD diagnosis/treatment solutions. This review paper will identify the best results and weaknesses of previous studies so that they can better deal with various challenges in their future works. In other words, this study identifies several strengths and weaknesses of previous data and methods. Additionally, current research benefits governments and patients. The accuracy, cost and speed of diagnosis/treatment of CAD are important issues for both of these groups. Thus, providing the best results and methods can be helpful for these individuals. We attempt to find answer to: (i) the best performing ML method, and (ii) find the best type of CAD data which will yield highest classification results. By knowing the answers to these questions, researchers, physicians, governments and patients can make decisions about using suitable ML methods.

2.1. Article selection strategy

We queried all datasets that are supported by Google scholar, including IEEE, Science Direct, Springer, DBLP Computer Science, ACM Digital Library, Hindawi, and PubMed. The search query was composed of the following terms:

(LAD OR LCX OR RCA OR CAD OR “coronary artery disease”) AND

Table 1
List of the selected studies.

Paper ID[a]	Reference[b]	Paper ID	Reference	Paper ID	Reference	Paper ID	Reference	Paper ID	Reference
P001	[49]	P003	[50]	P004	[51]	P005	[52]	P007	[53]
P008	[54]	P009	[55]	P010	[56]	P011	[57]	P012	[58]
P013	[59]	P015	[60]	P017	[61]	P018	[62]	P019	[63]
P021	[64]	P022	[65]	P023	[66]	P024	[67]	P026	[68]
P027	[69]	P028	[70]	P029	[71]	P030	[72]	P031	[73]
P032	[74]	P033	[75]	P034	[76]	P035	[77]	P037	[78]
P039	[79]	P040	[80]	P041	[81]	P042	[82]	P043	[83]
P044	[84]	P045	[85]	P046	[86]	P047	[87]	P050	[88]
P052	[89]	P054	[90]	P055	[91]	P056	[92]	P058	[93]
P059	[94]	P060	[95]	P061	[42]	P062	[45]	P065	[96]
P066	[97]	P067	[47]	P069	[98]	P070	[99]	P072	[100]
P073	[101]	P076	[102]	P080	[103]	P082	[104]	P083	[105]
P084	[106]	P085	[107]	P086	[108]	P087	[109]	P089	[110]
P091	[111]	P092	[48]	P095	[112]	P097	[113]	P098	[114]
P099	[115]	P100	[116]	P101	[117]	P102	[118]	P104	[119]
P106	[120]	P107	[121]	P110	[122]	P111	[123]	P112	[124]
P113	[125]	P114	[126]	P115	[46]	P116	[127]	P117	[128]
P118	[129]	P119	[130]	P120	[131]	P121	[132]	P122	[133]
P126	[134]	P127	[135]	P128	[136]	P129	[137]	P130	[138]
P131	[139]	P134	[140]	P135	[141]	P136	[142]	P137	[143]
P138	[144]	P139	[145]	P140	[146]	P142	[147]	P144	[148]
P146	[149]	P148	[150]	P149	[151]	P151	[152]	P152	[153]
P153	[46]	P154	[154]	P156	[155]	P158	[156]	P160	[157]
P161	[158]	P163	[159]	P164	[160]	P165	[161]	P166	[162]
P167	[163]	P168	[164]	P169	[165]	P170	[166]	P171	[167]
P174	[110]	P175	[168]	P230	[169]	P231	[170]	P232	[171]
P233	[172]	P234	[173]	P235	[174]				

[a] "Paper ID" is an identification number that is used in this research.
[b] "Reference" refers to the reference number of the corresponding paper ID. In this research, as we wanted to refer to other research in figures, instead of using the reference number, we used the paper ID because we want to avoid changing figures when the reference number changes.

(disease OR failure OR diagnosis OR prognosis OR treatment) AND ("machine learning" OR "data mining" OR "machine intelligent" OR classification OR clustering).

Returned articles were investigated by nine authors of this survey. Information about a research paper was added to the database if at least *five* coauthors considered it relevant to the scope of this review paper. In this way, 149 research articles related to ML-based CAD detection were identified. The full list of these papers can be found in Table 1. Fig. 2 provides detailed information about these publications.

The research on ML-based CAD detection commenced in 1992 (P076). There were not many papers until 2008, when there was a sharp jump in the number of papers. The growth in the number of papers continued until peaking in 2012 with 20 papers. The attention to this field was stable over the next years, followed by another peak in

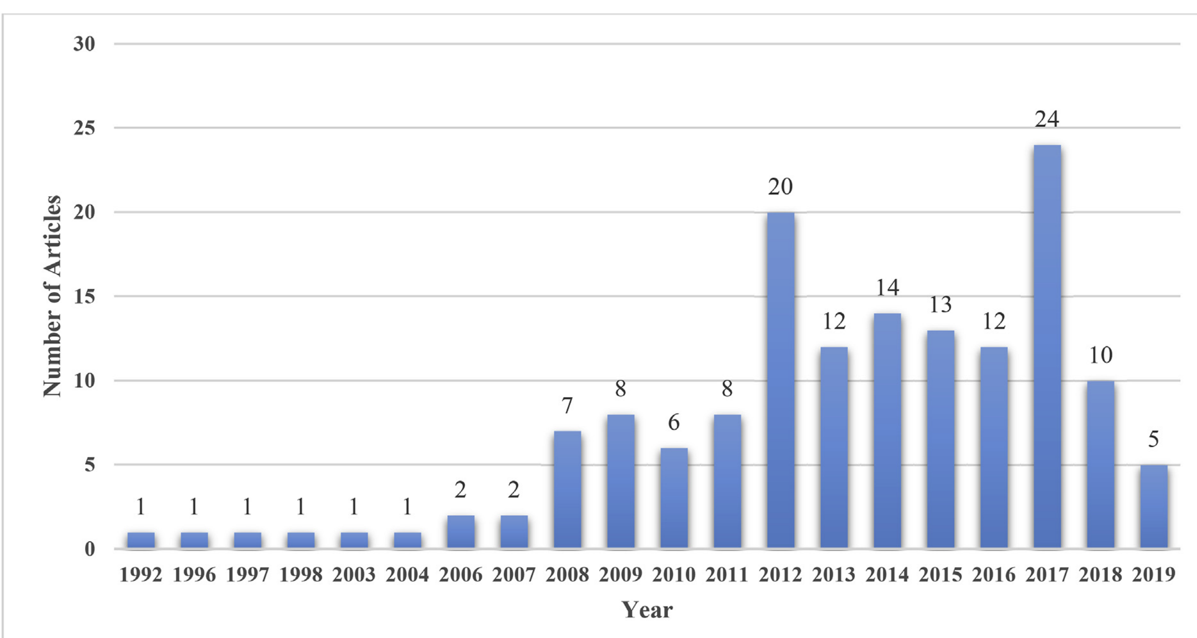


Fig. 2. Number of research papers published between 1992 and 2019. From 1992 to 2004, only one paper was published per year. The subject has attracted more attention since 2008 and reached a peak of 20 papers in 2012. However, the number of articles decreased between 2012 and 2013 and fluctuated between 2013 and 2016. In 2017, interest increased again and researchers published 24 papers. The turning points in CAD diagnosis occurred in 2012 and 2017.

2017. Further information about the most important papers in this field can be found in Supplementary Fig. S1 (see also the “Most cited papers” part of the Supplementary Information).

2.2. ML-based CAD diagnosis

ML-based CAD detection is a pure machine learning problem. While the model simplicity, interpretability, and computational burden are important factors, the doctors and practitioners are mainly concerned about the reliability and overall performance of the model in the detection of CAD. Several metrics, including accuracy, sensitivity, specificity, and f-score, have been reported in the relevant literature for model evaluation. Regardless of which metric is considered, the overall performance of the model depends on *two* key factors: (i) dataset, and (ii) ML pipeline. For the dataset, the analysis can be based on the source of the data (country), the sample size, and number of features. The pipeline can be investigated based on the ML and feature selection methods. The rest of Section 2 investigates these factors and conducts a multifaceted analysis of what ML-based CAD diagnosis has offered thus far.

2.3. Datasets

A total of 67 datasets are considered, which were collected in 18 countries/regions of 3 continents. We have listed only the datasets which were used in different articles to diagnose CAD using machine learning and data mining algorithms. These datasets differ significantly in terms of the numbers of samples and features. Fig. 3a displays a scatter plot of the numbers of features and samples; it is color-coded according to the data collection location. A scatter plot of the median values is shown in Fig. 3b. Marginal boxplots have been added to these two charts to represent data quartiles. The smallest dataset was collected in India and consists of only 20 samples and 9 features. The median sample size for all datasets is approximately 350. The largest dataset was also collected in India, which consists of more than 24,000 samples and 11 features. Hence, CAD detection datasets are small in size [175]. The limited number of samples in most datasets (data sparsity issue) could negatively impact the model performance. Furthermore, it renders the obtained results less generalizable as the samples used for model development may not cover all possible combinations of features.

The investigated articles have used various feature categories to detect CAD. For example, dataset IDs 38, 41, and 64 (Supplementary Table S2) only includes ECG features. Others, such as the UCI datasets, contain combinations of demographic, laboratory, symptom and examination, and echo features. Informative and important features for CAD detection can be shortlisted according to the accuracies reported for various datasets and ML methods. The most commonly used categories are demographic, symptom and examination, laboratory, ECG, fluoroscopy, and echo. It is much easier to collect data that belong to these categories as compared to angiography. Among these categories, collecting echo and fluoroscopy data is more difficult and requires special equipment.

Some datasets have many common features, and some have fewer features in common. Various demographic features, such as age and sex, were widely used in the datasets. However, various other features, such as genetic features, were only used in a single dataset. All collected datasets are listed in Supplementary Table S2.

The most commonly used datasets were uploaded to the UCI machine learning repository (datasets No. 1–7 in Supplementary Table S2). Among these, the Cleveland dataset, which was collected in the USA and consists of 303 samples, has been used more frequently than other datasets. The main problem of the first five datasets are having few features, old date of collection, and missing values. The two latest datasets, namely, (UCI) Z-Alizadeh Sani and Extension of Z-Alizadeh Sani, do not suffer from these problems. One limitation of most CAD datasets is that the data indicate only whether the patients have CAD or not. Only a few of them contain additional useful information, such as the stenosis severities of the LAD, LCX and RCA arteries. Therefore, few articles could detect arterial stenosis. The main difference between datasets 6 and 7 is this additional information. Extension of Z-Alizadeh Sani includes three additional features that represent the stenosis of the LAD, LCX and RCA arteries. Therefore, articles that use this dataset can not only analyze the stenosis of each artery separately but also determine the features that affect the stenosis of each artery.

2.4. Feature selection algorithms

Feature selection or dimension reduction refers to the process of selecting the most relevant variables or parameters from a list of features in the problem. Feature selection simplifies the model by reducing the number of required parameters (parsimonious model). It also

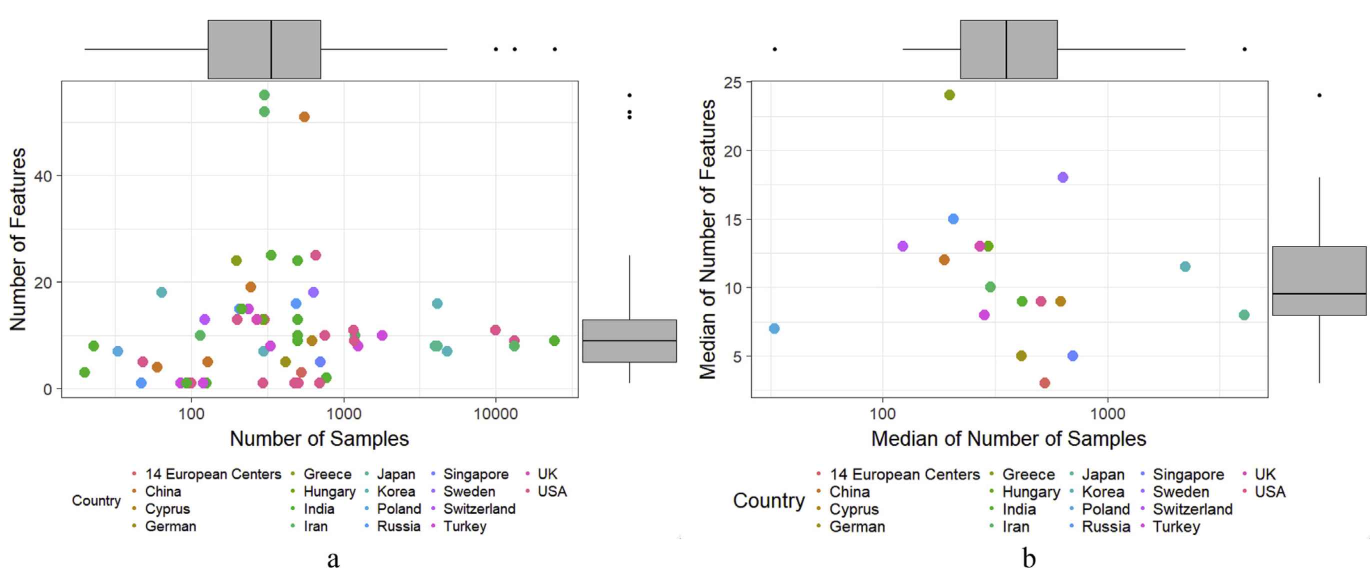


Fig. 3. Statistical review of research conducted in different countries: a) The numbers of features used versus the number of samples (the number of samples is shown on a logarithmic scale). b) The number of samples varies between 20 and + 24,000, with a median of approximately 350. The number of features also varies between 1 and 55, with a median of 9.

decreases the training time and prevents over-fitting [176]. Overall, it has the potential to increase the model's performance, especially for unseen samples. If the feature selection algorithms can find a suitable subset of features, this will help the ML algorithms converge faster and realize higher accuracy, which is mainly achieved by discarding redundant and irrelevant features without losing key information [176].

Feature selection algorithms use an evaluation measure to assign scores to features [177]. These scores are assigned according to the importance of features in determining the output labels. If the problem output is the diagnosis of a disease, it is possible to identify important factors for that disease, which will help physicians diagnose and cure diseases using more efficient methods. Therefore, most researchers used feature selection algorithms to identify important features for CAD.

Feature selection has a major impact on the model performance. It acts as a filter for muting irrelevant features and selecting a subset of more informative features for solving the problem [178]. A variety of feature selection methods have been applied for CAD detection. These methods include information gain, correlation, ANOVA, statistical tests (z, t, and F), principal component analysis, and the genetic algorithm [177]. Supplementary Fig. S6 presents the most popular feature selection methods in the diagnosis of CAD.

2.5. Model type

More than 90% of CAD detection papers have used supervised classification algorithms to process features and make decisions. Few articles reported applications of clustering methods for CAD diagnosis. Complete description of classification, clustering and other methods have been presented in the Supplementary Information (see the “Prediction and diagnosis of CAD” part of the Supplementary Information).

Fig. 4a shows the most frequently used ML methods for CAD detection. ANN, DTs, and SVM are the top *three* methods, which have been applied to almost all datasets that have been reported in the literature. Their popularity is due to their promising performance, ease of use, and low computational burden [179]. According to the fact that deep learning is the most powerful extension of ANN, it was investigated separately. The list of articles published using this technique for CAD diagnosis are shown in Table 2. In the second group are Naïve Bayes, KNN, and fuzzy rule-based systems, which have been applied mainly to well-known datasets. Researchers have also used other traditional classifiers such as random forest and logistic regression for CAD detection. These two recent classifiers are the most widely used classifiers for benchmarking in the field of supervised ML.

Fig. 4b shows the yearly usage frequencies of popular ML methods for CAD diagnosis. ANNs, DTs, and SVM have maintained their popularity over the years. ANNs have attracted more attention in this field. From 2004 to now, at least one research article has been published each year. In 2012, 2013 and 2017, 9, 10 and 10 research articles, respectively, used neural networks. DTs and SVM almost have the same usage frequencies as ANNs. There has been a noticeable increase in the number of papers that reported fuzzy rule-based models for CAD detection. Another more frequently used algorithm is Naïve Bayes. The usage of this algorithm has changed over the years. At first, its usage increased, especially in 2012. Then, it decreased slowly. In 2016, 2017 and 2018, only one research article was published each year. Although random forest performed well, few research studies used this algorithm.

2.6. Model performance

Model evaluation is one of the most important steps in the ML pipeline. The performance of a model can be measured via dozens of metrics. The measurements are often carried out on unseen samples during the training process. The most popular evaluation metric for CAD detection is accuracy. Relying purely on accuracy could be misleading, especially for highly imbalanced datasets [190]. Thus, accuracy should be always evaluated and interpreted in conjunction with other metrics, including sensitivity, specificity, and the area under the curve (AUC) [179]. Supplementary Fig. S2 shows the usage percentages of metrics in the investigated articles. Hereafter, we will investigate the model performance as measured by the accuracy metric.

Before going further, the importance of data splitting is highlighted. The common practice in the field of ML is to split the data into training, test and validation sets. The training set is used to adjust the model parameters via optimization of the loss function [179]. The validation set is applied to detect the onset of overfitting and to stop the training process. The performance metric is reported for the test set. The splitting of the original dataset into these three subsets could have a significant impact on the overall performance of the model. There are numerous methods for data splitting. Supplementary Fig. S3 summarizes the settings of the training, test, and validation sets that are used in the relevant literature for CAD.

For each year, Fig. 5a shows the maximum accuracies of algorithms that were commonly used to detect CAD. As the most famous datasets are UCI datasets, we provided additional details about the maximum accuracy on them in the Supplementary Information (see the “UCI heart dataset accuracy analysis” part of the Supplementary Information). Fig. 5b compares several of these algorithms. This figure facilitates pairwise comparison of algorithms. For example, Naïve Bayes has not been able to outperform C4.5. Additionally, in all papers except one, Naïve Bayes performed worse than SVM. The performance of ANN in all papers except one was worse than those of SVM and fuzzy rule-based system (FRBS). Toward the end of the x-axis, because many comparisons have been shown already, the number of comparisons is reduced. Supplementary Fig. S4 shows the maximum accuracies that are achieved by various classification algorithms in diagnosing CAD.

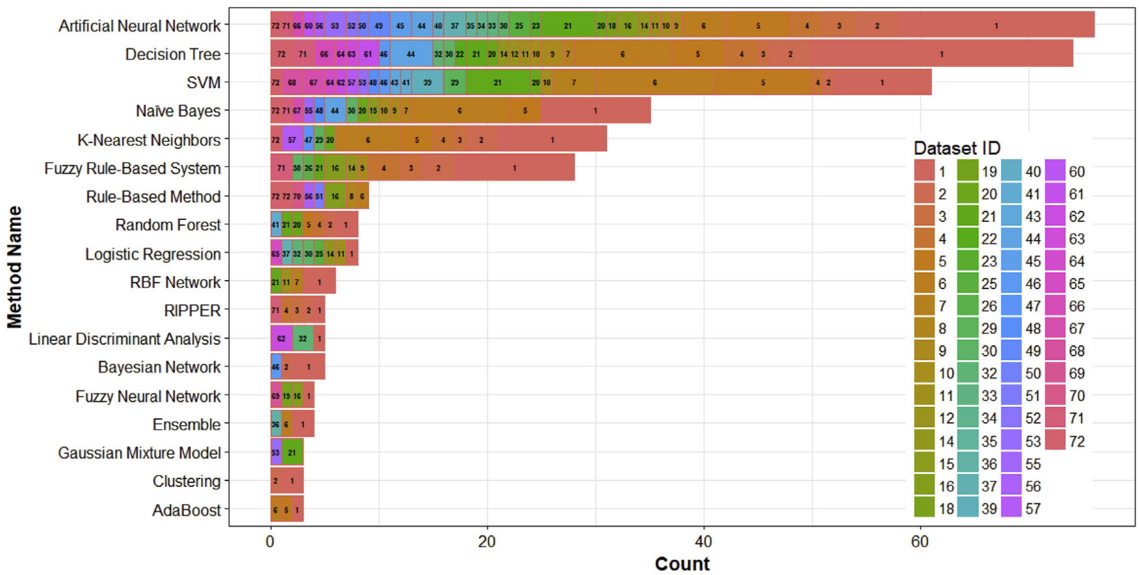
More information about Fig. 5b and how the methods are compared is provided in Supplementary Table S3 [191]. This table quantitatively compares the methods that are applied in the same paper in terms of accuracy.

2.7. LAD, LCX and RCA investigation

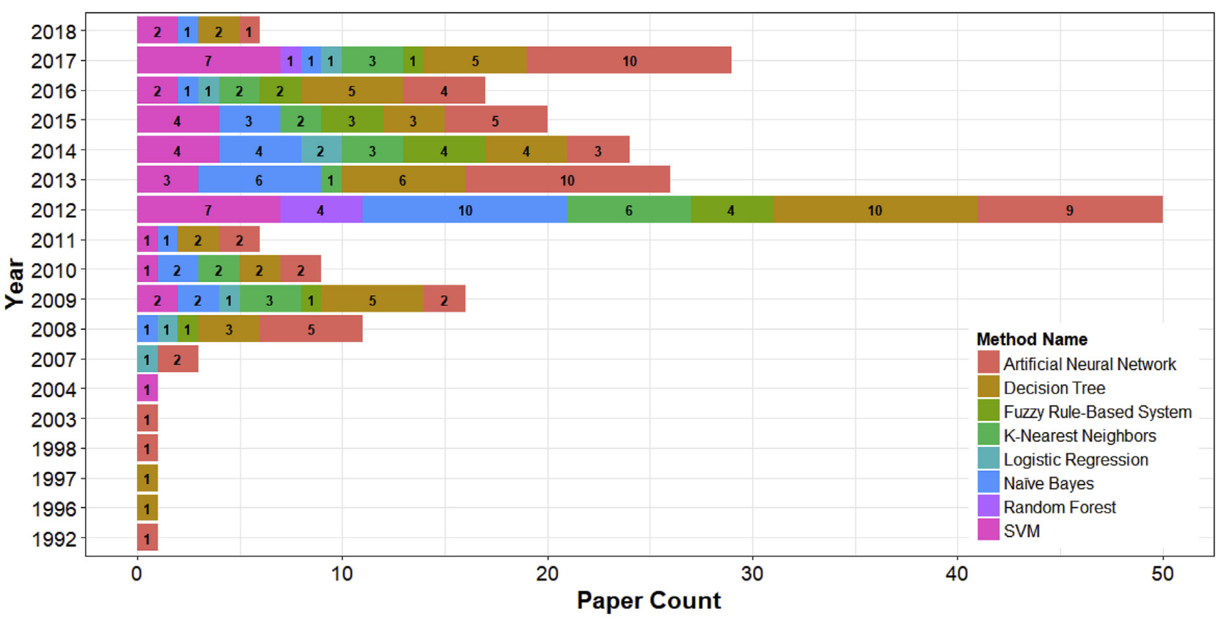
Supplementary Fig. S5 illustrates the studies on stenosis of the LAD, LCX and RCA arteries. The arterial stenosis percentage, which is valuable information in the evaluation of a method for CAD diagnosis, has not yet been considered. This percentage indicates the amount of stenosis, rather than only whether an artery is stenotic or not. Therefore, a promising direction for future research is to collect additional datasets and to consider every artery separately to determine the level of arterial stenosis. The best method for diagnosing CAD is angiography, in which three main outputs are determined: 1. whether the patient has CAD or not, 2. which artery is stenotic, and 3. the percentage of the stenosis. The first output is already well considered in the literature. In contrast, the second output is only studied in 5 articles, with the highest accuracy of 86.1% for LAD stenosis diagnosis (P065). Surprisingly, there is no study that reports on the third output. Hence, reducing or even eliminating the use of angiography due to its costs and side effects requires extensive research and work on the second and third outputs of angiography.

3. Discussion and future work

In this section, we not only listed the papers with the best reported performance but also investigated the impact of feature categories on the performance of ML algorithms. Meanwhile, this requires further investigation to determine which classification algorithms work better for a specific feature category. Feature category is a branch of features that have been extracted using a specific method. For example, echo is a category that features such as ejection fraction, regional wall motion abnormality, and valvular heart disease can be extracted from. The most commonly used categories are demographic, symptom and



a



b

Fig. 4. Most commonly used machine learning and data mining algorithms for CAD detection. a) Artificial neural network, decision tree, and SVM are the most preferred methods in this field. Naïve Bayes, KNN and fuzzy rule-based system methods are the next most frequently used learning techniques in articles. This figure also illustrates the datasets that have been used by each classifier. Most of these methods have been applied to the UCI datasets with dataset IDs 1 to 7. Dataset ID 1, which is named Cleveland, is the most popular one. This figure also shows the methods that have been used less frequently than the others and the methods that have not been used. These methods should be considered for future research in the field. **b)** The method usage changes over the years. The largest changes occurred in the years 2012 and 2017. Artificial neural networks and decision tree were implemented for the first time in 1992 and 1996, respectively. Logistic regression was first applied to CAD data in 2007. Naïve Bayes and the fuzzy rule-based system were implemented in 2008. KNN was used for the first time in 2009. Overall, the mix of methods has remained almost the same, although the application of computational-intelligence-based methods, including neural networks and fuzzy logic systems, has increased in recent years. This result is in full agreement with the recent advances in these fields [5].

examination, laboratory, ECG, fluoroscopy, and echo, respectively. The research papers with the best reported performance have been summarized in Table 3. In this table, the papers with a reported accuracy more than 98% were sorted by their performance in descending order. We found that, papers with best performances used ANN and SVM as their classifiers. It is not surprising because these two classifiers are among the best ones in other fields too. This may be due to the use of nonlinear kernel functions. In Table 4, most used categories and their properties have been shown.

3.1. Challenges and disadvantages of using ML algorithms

Although ML techniques have many advantages, they are not perfect methods. The following factors limit their abilities in some directions [179].

1. According to no-free-lunch theorem [192], different ML algorithms are suitable for their own particular problem. One algorithm may work well on a specific dataset while it cannot show a good performance on some others. So, selecting a suitable algorithm for a specific dataset is a big challenge in bioinformatics. Consequently,

Table 2
The list of articles published using deep learning method for CAD diagnosis.

Authors	Year	Techniques	Number of participants	Sampling	Conclusion	Data
Shen et al. [180]	2019	3D fully convolutional network (FCN)	70	Train = 70% Test = 30%	Accuracy = 90.05% Sensitivity = NR[a] Specificity = NR	Coronary computed tomography angiography (CTA)
Acharya et al. [181]	2017	11- layer deep Convolutional Neural Network (CNN)	200	10-fold cross-validation	Accuracy = 95.22% Sensitivity = 95 .49% Specificity = 94 .19%	ECG signals
Betancur et al. [182]	2018	6-layer deep CNN	1638	10-fold cross-validation	Sensitivity = 82.3% Specificity = 68%	Myocardial perfusion imaging (MPI)
Betancur et al. [183]	2018	Deep CNN	1160	Four-fold cross-validation	Sensitivity = 84.8% Accuracy = NR Specificity = NR	SPECT MPI
Abdolmanafi et al. [184]	2018	5-layer AlexNet architecture	33	10-fold cross-validation	Accuracy = 94% Sensitivity = 90% Specificity = 99%	Optical Coherence Tomography (OCT)
Yeri et al. [185]	2018	6-layer deep CNN	351	10-fold cross-validation	Accuracy = 78%	CT angiography
Rubin et al. [186]	2017	6-layer deep CNN	90000	Train = 80% Test = 20%	Sensitivity = 73% Specificity = 95% Accuracy = 84%	Heart sound
Hamersvelt et al. [187]	2018	6-layer deep CNN	126	10-fold cross-validation	Accuracy = 71.1% Sensitivity = 84.6% Specificity = 48.4%	Resting coronary CT angiography (CCTA)
Sofian et al. [188]	2018	34 layers ResNet101 architecture	10	10-fold Cross-Validation	Accuracy = 99.49% Sensitivity = 99.25% Specificity = 99.57%	Intravascular ultrasound (IVUS) image
Zreik et al. [189]	2017	8-layer deep CNN	166	10-fold cross-validation	Accuracy = 71% Sensitivity = 70% Specificity = 71%	CCTA
Tan et al. [137]	2018	8-layer deep CNN	47	10-fold cross-validation	Accuracy = 99.85% Sensitivity = 99.85% Specificity = 99.84%	ECG signals
Acharya et al. [94]	2017	11-layer deep CNN	47	10-fold cross-validation	Accuracy = 95.11% Sensitivity = 91.13% Specificity = 95.88%	ECG signals
Acharya et al. [84]	2019	11-layer deep CNN	40	10-fold cross-validation	Accuracy = 98.97% Sensitivity = 98.87% Specificity = 99.01%	ECG signals
Allahverdi et al. [163]	2016	Deep belief network (DBN)	85	10-fold cross-validation	Accuracy = 98.05% Sensitivity = 96.02% Specificity = 98.88%	ECG signals

[a] NR means Not Reported.

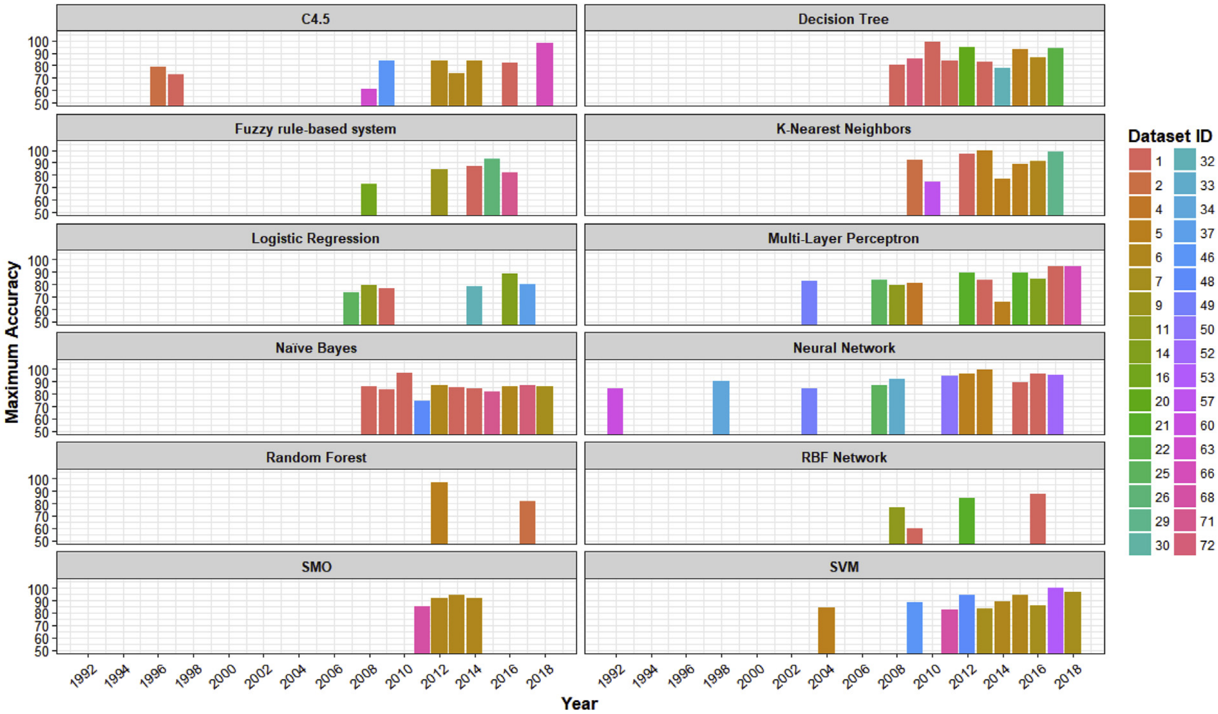
selecting good feature selection or classification algorithms is also a big challenge in this field.
2. ML algorithms commonly need massive datasets to be trained. These datasets must be inclusive and unbiased with high quality. Datasets also need time to be collected.
3. ML algorithms need time to be trained and tested enough to be able to generate results with high confidence. These algorithms need a lot of resources and equipment.
4. ML algorithms face the verification problem. It is difficult to prove that the prediction made by them work correctly for all scenarios.
5. Correct interpretation of the generated results by ML algorithms is another challenge that we are faced with.
6. Another disadvantage of ML algorithms is their high error-susceptibility. If they are trained with biased or incorrect data, they end up with imprecise outputs. This may lead to a chain of errors that mislead the treatment methods. When these errors get noticed, it takes some times to diagnose the source of these errors and even need more time to correct them [193].

3.2. Perspectives on future work

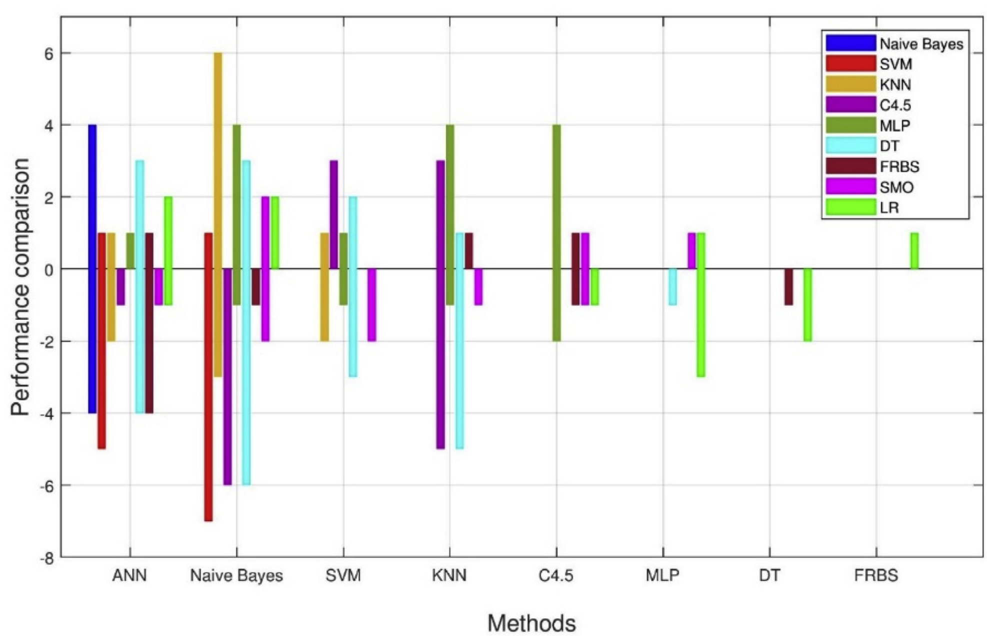
Even though various ML methods and data processing techniques have been proposed for CAD detection, there are few avenues for improvement that require deep attention from researchers, professionals, doctors and, ultimately, governments. This section discusses potential directions for further research. There are three major subfields that can be considered as future works. At first, improvement of the mechanism used for collection and management of datasets. Then, finding suitable algorithms for analyzing the collected data and finally investigating each major coronary artery, separately. These subfields are described in following paragraphs.

There are multiple issues with the datasets used for ML-based CAD analysis. They are summarized as follows:

- Lack of information for many regions: In this study 149 papers have applied ML methods to datasets that were collected in a few countries. There is no published data for most countries in Europe, Africa, Australia, and South America. This lack of information is important as regional and racial differences may affect the way CAD is detected and treated. Thus, we recommend collecting CAD data and constructing databases from various continents and countries.
- Limited features: Most of investigated datasets have limited number of features. This severely limits the final results since the numbers of both samples and features can affect the performance of ML techniques. Hence, we suggest constructing CAD databases with more features.
- Small sample sizes: The larger the number of samples is, the more significant is the statistical results and more conclusive will be the resulting decisions. The median sample size for the CAD datasets that were investigated in this review is less than 500. Thus, the



a



b

Fig. 5. Comparing of the most frequently used machine learning methods in CAD diagnosis. a) This figure shows the maximum accuracy of each method per year on the corresponding dataset. Most of the top accuracies are obtained on UCI datasets (numbers 1–7). **b)** This chart compares 10 important algorithms that have the most applications in CAD diagnosis in terms of accuracy. For example, compared to Naïve Bayes, which is shown in dark blue, the ANN algorithm, which is located on the horizontal axis, shows superior performance in four articles. It shows poorer performance in the other four articles. Meanwhile, the SVM algorithm, which is shown in red, substantially outperforms the ANN and Naïve Bayes algorithms. Additionally, the C4.5 algorithm, which is shown in dark pink, substantially outperforms the KNN and Naïve Bayes algorithms. It is concluded that algorithms that are located on the horizontal axis for which the bar charts are mostly below the zero level of the vertical axis had poor performance. For example, the ANN classification algorithm has the weakest performance among all methods except LR and MLP, as the ANN's bar charts, except the green and dark green ones, are below the zero level of the vertical axis.

generalization power of the developed models is questionable. To ensure reliability and trustworthiness, a model should be developed and tested using at least one million samples [194].

- Another problem with previous studies is the way the data were collected. Since the datasets differ in terms of the number of samples and features, it is not easy to compare ML techniques in terms of performance. In other words, the results obtained in various studies are comparable only if the data are the same. However, a highly

Table 3
The summary of research papers with accuracy more than 98%.

Authors	Year	Techniques	Number of participants	Number of features	Sampling	Performance	Data
Jabbar et al. [73]	2013	Artificial Neural Network + PCA	23	8	10-fold cross-validation	Accuracy = 100% Sensitivity = 100% Specificity = 100%	Dataset ID 44
Jabbar et al. [91]	2013	KNN	23	8	10-fold cross-validation	Accuracy = 100% Sensitivity = 100% Specificity = 100%	Dataset ID 44
Kumar et al. [148]	2016	LS-SVM with RBF and Morlet wavelet kernel	20	4	10-fold cross-validation	Accuracy = 100% Sensitivity = 100% Specificity = 100%	Heart Rate Variability (HRV) signals
Acharya et al. [172]	2013	Gaussian Mixture Model (GMM)	800	559	3-fold cross-validation	Accuracy = 100% Sensitivity = 100% Specificity = 100%	Grayscale features [172]
Ikeda et al. [130]	2013	SVM	501	NR	NR	Accuracy = 100% Sensitivity = 100% Specificity = 100%	Grayscale features
Tan et al. [137]	2018	8-layer deep convolutional neural network (CNN)	47	NR	10-fold cross-validation	Accuracy = 99.85% Sensitivity = 99.85% Specificity = 99.84%	ECG signals
Patidar et al. [138]	2015	LS-SVM	20	3	3-fold cross-validation	Accuracy = 99.7% Sensitivity = 99.63% Specificity = 99:81%	ECG signals
Jabbar et al. [91]	2013	Artificial Neural Network	270	13	10-fold cross-validation	Accuracy = 99.62% Sensitivity = NR Specificity = NR	Dataset ID 5
Kumar et al. [110]	2017	LS-SVM	207	15	10-fold cross-validation	Accuracy = 99.6% Sensitivity = 99.57% Specificity = 99.61%	ECG signals
Acharya et al. [108]	2017	KNN	222	495	10-fold cross-validation	Accuracy = 99.55% Sensitivity = 99.93% Specificity = 99.24%	ECG signals
Sharma et al. [174]	2019	Gaussian support vector machine (GSVM)	254	16	10-fold cross-validation	Accuracy = 99.53% Sensitivity = 98.64% Specificity = 99.70%	ECG signals
Anbarasi et al. [116]	2010	Decision Tree	920	13	10-fold cross-validation	Accuracy = 99.2% Sensitivity = NR Specificity = NR	Dataset IDs 1-4
Acharya et al. [112]	2017	Decision Tree	47	13	10-fold cross-validation	Accuracy = 98.99% Sensitivity = 97.75% Specificity = 99.39%	ECG signals
Acharya et al. [84]	2019	11-Layer Deep Convolutional Neural Network	40	NR	10-fold cross-validation	Accuracy = 98.97% Sensitivity = 98.87% Specificity = 99.01%	ECG signals
Rajkumar et al. [145]	2017	Fuzzy Neural network	768	13	10-fold cross-validation	Accuracy = 98.88% Sensitivity = 99.61% Specificity = 71.42%	Dataset IDs 1,2, 5
Sridhar et al. [160]	2016	KNN	47	20	10-fold cross-validation	Accuracy = 98.67% Sensitivity = 95.02% Specificity = 99.20%	ECG signals
Acharya et al. [95]	2017	KNN	207	7	10-fold cross-validation	Accuracy = 98.5% Sensitivity = 99.7% Specificity = 98.5%	ECG signals
Acharya et al. [108]	2017	Decision Tree	170	495	10-fold cross-validation	Accuracy = 98.39% Sensitivity = 99.74% Specificity = 95.79%	ECG signals
Acharya et al. [112]	2017	KNN	47	31	10-fold cross-validation	Accuracy = 98.17% Sensitivity = 94.57% Specificity = 99.34%	ECG signals
Jabbar et al. [73]	2013	Artificial Neural Network + PCA	270	13	10-fold cross-validation	Accuracy = 98.14% Sensitivity = NR Specificity = NR	Dataset ID 5
Allahverdi et al. [163]	2016	Deep Belief Networks (DBN)	47	9	10-fold cross-validation	Accuracy = 98.05% Sensitivity = 96.02% Specificity = 98.88%	ECG signals

specialized medical perspective is required and only CAD specialists can propose important factors. To this end, closer and more effective communication between specialists in medical and ML teams should be considered.

ML algorithms can also be improved in some directions. For example:

- Application of advanced feature section techniques to identify the most informative features. This will lead to the development of parsimonious models that are resistant to noise and uncertainties in the data [195].
- Application of evolutionary computation algorithms for model development and feature engineering. These algorithms avoid becoming trapped in local minima. Hence, they can improve the

Table 4
Properties of feature categories used in more than three references.

Feature category	Number of papers used this category	Best Classification algorithm	Reference *	Maximum Accuracy %	Sensitivity % **	Specificity % **
Demographic	79	Artificial Neural Network	P031	100	100	100
Symptom and Examination	69	Artificial Neural Network	P031	100	100	100
Laboratory	68	Decision Tree	P100	99.2	N.R.	N. R.
ECG	58	Least Squares-Support Vector Machine	P144	100	100	100
Fluoroscopy	38	Decision Tree	P100	99.2	N.R.	N. R.
Echo	9	Support Vector Machine	P005	96.4	100	88.1
Heart rate signals	3	Least Squares-Support Vector Machine	P144	100	100	100

* Reference paper of best classification algorithm.
** Sensitivity and specificity of the method with maximum accuracy.

performance of ML-based CAD detection [196].

- The ML research field has been revolutionized in recent years by the introduction of deep learning (DL) algorithms [94,137,197]. DL has achieved superhuman performance on various tasks, including driving cars, games, robotics, disease diagnosis, and computer vision [20]. These advances have ignited a boom in DL applications for solving complex problems in various fields of science and engineering. No doubt, their application for CAD detection could remarkably improve the classification results. However, DL models require massive datasets, which are not available for CAD case studies.
- Most ML-based CAD diagnosis studies have focused on developing individual models. Ensemble-based learning (boosting and bagging) techniques can be applied to remarkably improve the performance of individually weak models [191].

Angiography is the standard approach for the diagnosis of CAD. Most studies in the literature used this approach to determine whether the patient has CAD or not. However, angiography can be used to determine which artery is stenotic and the percentage of stenosis. Only five studies reported the application of ML in determining which artery is stenotic and the classification results for these studies are poor. As we discussed above, the highest accuracy in identifying LAD stenosis was 86.10% [96]. Hence, improvement of the algorithms in this area is required to realize better prediction performance. Therefore, additional research is strongly recommended. Moreover, there is no study in which the category three output (percentage of stenosis) is determined via ML methods.

4. Conclusions

In this review paper, we have investigated several studies on the detection of CAD using various ML techniques. To this end, we have conducted a comprehensive search using various search engines and databases. According to our review, the most important studies on the diagnosis of CAD via ML algorithms were conducted from 2012 to 2017. Additionally, the results demonstrated that ANNs, DTs, SVMs, Naïve Bayes, and KNN are the most widely used algorithms for CAD detection. Due to inherent differences among datasets, inconsistent performances have been reported for different datasets using similar ML algorithms. The reported results indicate that KNN, SVM, and ANN have achieved the highest accuracies for most of the CAD datasets. Lastly, there are several common features among the datasets. These include age, blood pressure, chest pain type, sex, total cholesterol, ST depression, hypertension, and maximum heart rate. Despite the progress that has been made in recent years, there remain key shortcomings in ML-based detection of CAD that must be addressed in upcoming years.

Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its Supplementary Information.

Conflicts of interest

The authors declare no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at h.ttps://doi.org/10.1016/j.compbiomed.2019.103346

References

[1] S.J. Russell, P. Norvig, Artificial Intelligence: a Modern Approach, Pearson Education Limited, Malaysia, 2016.
[2] M. Waldrop, No drivers required, Nature 518 (2015) 20.
[3] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, D. Hassabis, Human-level control through deep reinforcement learning, Nature 518 (2015) 529.
[4] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis, Mastering the game of Go without human knowledge, Nature 550 (2017) 1–6.
[5] M. Wainberg, D. Merico, A. Delong, B.J. Frey, Deep learning in biomedicine, Nat. Biotechnol. 36 (2018) 829.
[6] P. Eulenberg, N. Köhler, T. Blasi, A. Filby, A.E. Carpenter, P. Rees, F.J. Theis, F.A. Wolf, Reconstructing cell cycle and disease progression using deep learning, Nat. Commun. 8 (2017) 1–6.
[7] J. De Fauw, J.R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O'Donoghue, D. Visentin, G. van den Driessche, B. Lakshminarayanan, C. Meyer, F. Mackinder, S. Bouton, K. Ayoub, R. Chopra, D. King, A. Karthikesalingam, C.O. Hughes, R. Raine, J. Hughes, D.A. Sim, C. Egan, A. Tufail, H. Montgomery, D. Hassabis, G. Rees, T. Back, P.T. Khaw, M. Suleyman, J. Cornebise, P.A. Keane, O. Ronneberger, Clinically applicable deep learning for diagnosis and referral in retinal disease, Nat. Med. 24 (2018) 1342–1350.
[8] S.L.F. Walsh, L. Calandriello, M. Silva, N. Sverzellati, Deep learning for classifying fibrotic lung disease on high-resolution computed tomography: a case-cohort study, Lancet Respir. Med. (2018) 1–9.
[9] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, I. Chouvarda, Machine learning and data mining methods in diabetes research, Comput. Struct. Biotechnol. J. 15 (2017) 104–116.
[10] S.K. Shrivastava, S.S. Gharde, Support vector machine for handwritten Devanagari numeral recognition, Int. J. Comput. Appl. 7 (2010) 9–14.
[11] N. Amruthnath, T. Gupta, Fault class prediction in unsupervised learning using model-based clustering approach, International Conference on Information and Computer Technologies (ICICT), 2018, pp. 5–12.
[12] H. Hassanzadeh, M. Keyvanpour, A Machine Learning Based Analytical Framework for Semantic Annotation Requirements, (2011) arXiv preprint arXiv:1104.4950.
[13] T. Ni, X. Gu, J. Wang, Y. Zheng, H. Wang, Scalable transfer support vector machine with group probabilities, Neurocomputing 273 (2018) 570–582.
[14] X. Zhu, Z. Huang, H. Tao Shen, J. Cheng, C. Xu, Dimensionality reduction by mixed kernel canonical correlation analysis, Pattern Recogn. 45 (2012) 3003–3016.

[15] G. Schmidberger, Tree-based Density Estimation: Algorithms and Applications, The University of Waikato, 2009.
[16] Y. Zhang, Q. Yang, A Survey on Multi-Task Learning, (2017) arXiv preprint arXiv:.08114.
[17] H. Brink, J.W. Richards, M. Fetherolf, Real-world Machine Learning, (2017) Manning.
[18] A. Géron, Hands-on Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly Media, Inc, 2017.
[19] H. Yan, Q. Ye, T.a. Zhang, D.-J. Yu, X. Yuan, Y. Xu, L. Fu, Least squares twin bounded support vector machines based on L1-norm distance metric for classification, Pattern Recogn. 74 (2018) 434–447.
[20] J. Schmidhuber, Deep learning in neural networks: an overview, Neural Network. 61 (2015) 85–117.
[21] M. Jaworski, P. Duda, L. Rutkowski, New splitting criteria for decision trees in stationary data streams, IEEE Trans. Neural Network. Learn Syst. 29 (2018) 2516–2529.
[22] C.-z. Gao, Q. Cheng, P. He, W. Susilo, J. Li, Privacy-preserving Naive Bayes classifiers secure against the substitution-then-comparison attack, Inf. Sci. 444 (2018) 72–88.
[23] S. Zhang, D. Cheng, Z. Deng, M. Zong, X. Deng, A novel kNN algorithm with data-driven k parameter computation, Pattern Recogn. Lett. 109 (2018) 44–54.
[24] D. Cheng, Y. Shi, T. Lin, B. Gwee, K. Toh, Hybrid K-means clustering and support vector machine method for via and metal line detections in delayered IC images, IEEE Trans. Circuits Syst. II: Express Briefs 4 (2018) 1.
[25] M. Abdar, M. Zomorodi-Moghadam, R. Das, I.H. Ting, Performance analysis of classification algorithms on early detection of liver disease, Expert Syst. Appl. 67 (2017) 239–251.
[26] M. Abdar, N.Y. Yen, J.C.-S. Hung, Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision trees, J. Med. Biol. Eng. 10 (2017) 1–13.
[27] P. Pławiak, Novel genetic ensembles of classifiers applied to myocardium dysfunction recognition based on ECG signals, Swarm Evol. Comput. 39 (2018) 192–208.
[28] P. Pławiak, Novel methodology of cardiac health recognition based on ECG signals and evolutionary-neural system, Expert Syst. Appl. 92 (2018) 334–349.
[29] E. Abdulhay, N. Arunkumar, K. Narasimhan, E. Vellaiappan, V. Venkatraman, Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease, Future Gener. Comput. Syst. 83 (2018) 366–373.
[30] M. Abdar, M. Zomorodi-Moghadam, Impact of patients' gender on Parkinson's disease using classification algorithms, J. AI Data Min. 6 (2018) 277–285.
[31] F. Khozeimeh, R. Alizadehsani, M. Roshanzamir, A. Khosravi, P. Layegh, S. Nahavandi, An expert system for selecting wart treatment method, Comput. Biol. Med. 81 (2017) 167–175.
[32] F. Khozeimeh, F. Jabbari Azad, Y. Mahboubi Oskouei, M. Jafari, S. Tehranian, R. Alizadehsani, P. Layegh, Intralesional immunotherapy compared to cryotherapy in the treatment of warts, 56 (2017) 474–478.
[33] R. Alizadehsani, M. Abdar, S.M.J. Jalali, M. Roshanzamir, A. Khosravi, S. Nahavandi, Comparing the performance of feature selection algorithms for wart treatment selection, Proc. Int. Workshop Future Technol. (2018) 6–18.
[34] C. Castaneda, K. Nalley, C. Mannion, P. Bhattacharyya, P. Blake, A. Pecora, A. Goy, K.S. Suh, Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine, J. Clin. Bioinform. 5 (2015) 1–16.
[35] L.A. Kelly-Hope, J. Cano, M.C. Stanton, M.J. Bockarie, D.H. Molyneux, Vectors, Innovative tools for assessing risks for severe adverse events in areas of overlapping Loa loa and other filarial distributions: the application of micro-stratification mapping, Parasite 7 (2014) 1–7.
[36] T. Laksanasopin, T.W. Guo, S. Nayak, A.A. Sridhara, S. Xie, O.O. Olowookere, P. Cadinu, F. Meng, N.H. Chee, J. Kim, C.D. Chin, E. Munyazesa, P. Mugwaneza, A.J. Rai, V. Mugisha, A.R. Castro, D. Steinmiller, V. Linder, J.E. Justman, S. Nsanzimana, S.K. Sia, A smartphone dongle for diagnosis of infectious diseases at the point of care, Sci. Transl. Med. 7 (2015) 273re271.
[37] E. Long, H. Lin, Z. Liu, X. Wu, L. Wang, J. Jiang, Y. An, Z. Lin, X. Li, J. Chen, J. Li, Q. Cao, D. Wang, X. Liu, W. Chen, Y. Liu, An artificial intelligence platform for the multihospital collaborative management of congenital cataracts, Nat. Biomed. Eng. 1 (2017) 0024.
[38] Cardiovascular diseases (CVDs), http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds accessed on 30/9/2018.
[39] D.P. Zipes, P. Libby, R.O. Bonow, D.L. Mann, G.F. Tomaselli, Braunwald's Heart Disease E-Book: A Textbook of Cardiovascular Medicine, Elsevier Health Sciences, 2018.
[40] http://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death Last access:1/10/2018.
[41] https://www.worldlifeexpectancy.com/cause-of-death/coronary-heart-disease/by-country/.Last.access 1/10/2018.
[42] R. Alizadehsani, J. Habibi, M.J. Hosseini, H. Mashayekhi, R. Boghrati, A. Ghandeharioun, B. Bahadorian, Z.A. Sani, A data mining approach for diagnosis of coronary artery disease, Comput. Methods Progr. Biomed. 111 (2013) 52–61.
[43] J. Nahar, T. Imam, K.S. Tickle, Y.-P.P. Chen, Computational intelligence for heart disease diagnosis: a medical knowledge driven approach, Expert Syst. Appl. 40 (2013) 96–104.
[44] J. Nahar, T. Imam, K.S. Tickle, Y.-P.P. Chen, Association rule mining to detect factors which contribute to heart disease in males and females, Expert Syst. Appl. 40 (2013) 1086–1093.
[45] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, A.A. Yarifard, Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm, Comput. Methods Progr. Biomed. 141 (2017) 19–26.
[46] M.G. Tsipouras, T.P. Exarchos, D.I. Fotiadis, A.P. Kotsia, K.V. Vakalis, K.K. Naka, L.K. Michalis, Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling, IEEE Trans. Inf. Technol. Biomed. 12 (2008) 447–458.
[47] P.K. Anooj, Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules, J. King Saud Univ. Comput. Inf. Sci. 24 (2012) 27–40.
[48] D. Pal, K.M. Mandana, S. Pal, D. Sarkar, C. Chakraborty, Fuzzy expert system approach for coronary artery disease screening using clinical parameters, Knowl. Based Syst. 36 (2012) 162–174.
[49] Y. Xuesong, C. Qiang, C. Yuquan, Noninvasive detection of coronary artery disease based on heart sounds, International Conference of the IEEE Engineering in Medicine and Biology Society, 3 1998, pp. 1546–1548.
[50] N. Bhatla, K. Jyoti, An analysis of heart disease prediction using different data mining techniques, Int. J. Eng. 1 (2012) 1–4.
[51] A. Kampouraki, G. Manis, C. Nikou, Heartbeat time series classification with support vector machines, IEEE Trans. Inf. Technol. Biomed. 13 (2009) 512–518.
[52] R. Alizadehsani, M.J. Hosseini, A. Khosravi, F. Khozeimeh, M. Roshanzamir, N. Sarrafzadegan, S. Nahavandi, Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries, Comput. Methods Progr. Biomed. 162 (2018) 119–127.
[53] M. Motwani, D. Dey, D.S. Berman, G. Germano, S. Achenbach, M.H. Al-Mallah, D. Andreini, M.J. Budoff, F. Cademartiri, T.Q. Callister, H.-J. Chang, K. Chinnaiyan, B.J.W. Chow, R.C. Cury, A. Delago, M. Gomez, H. Gransar, M. Hadamitzky, J. Hausleiter, N. Hindoyan, G. Feuchtner, P.A. Kaufmann, Y.-J. Kim, J. Leipsic, F.Y. Lin, E. Maffei, H. Marques, G. Pontone, G. Raff, R. Rubinshtein, L.J. Shaw, J. Stehli, T.C. Villines, A. Dunning, J.K. Min, P.J. Slomka, Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis, Eur. Heart J. 38 (2017) 500–507.
[54] A.F. Otoom, E.E. Abdallah, Y. Kilani, A. Kefaye, M. Ashour, Effective diagnosis and monitoring of heart disease, Heart 9 (2015) 143–156.
[55] M.C. Çolak, C. Çolak, H. Kocatürk, S. Sagiroglu, İ. Barutçu, Predicting coronary artery disease using different artificial neural network models/Koroner arter hastaliginin degisik yapay sinir agi modelleri ile tahmini, Anadulu Kardiyoloji Dergisi: AKD 8 (2008) 249–252.
[56] Q.A. Rahman, L.G. Tereshchenko, M. Kongkatong, T. Abraham, M.R. Abraham, H. Shatkay, Identifying hypertrophic cardiomyopathy patients by classifying individual heartbeats from 12-lead ECG signals, proceedings, IEEE Int. Conf. Bioinf. Biomed. (2014) 224–229.
[57] S. Rosenberg, M.R. Elashoff, P. Beineke, et al., MUlticenter validation of the diagnostic accuracy of a blood-based gene expression test for assessing obstructive coronary artery disease in nondiabetic patients, Ann. Intern. Med. 153 (2010) 425–434.
[58] M. Green, J. Björk, J. Forberg, U. Ekelund, L. Edenbrandt, M. Ohlsson, Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room, Artif. Intell. Med. 38 (2006) 305–318.
[59] J.R. Quinlan, Improved use of continuous attributes in C4. 5, J. Artif. Intell. Res. 4 (1996) 77–90.
[60] B. Kolukisa, H. Hacilar, G. Goy, M. Kus, B. Bakir-Gungor, A. Aral, V.C. Gungor, Evaluation of classification algorithms, linear discriminant analysis and a new hybrid feature selection methodology for the diagnosis of coronary artery disease, IEEE Int. Conf. Big Data (Big Data) (2018) 2232–2238.
[61] Z.-W. Chen, Y.-H. Chen, J.-Y. Qian, J.-Y. Ma, J.-B. Ge, Validation of a novel clinical prediction score for severe coronary artery diseases before elective coronary angiography, PLoS One 9 (2014) e94493.
[62] R. Rajkumar, K. Anandakumar, A. Bharathi, Coronary artery disease (CAD) prediction and classification-a survey, Breast Canc. 90 (2006) 94.35.
[63] V. Mahesh, A. Kandaswamy, C. Vimal, B. Sathish, ECG arrhythmia classification based on logistic model tree, J. Biomed. Sci. Eng. 2 (2009) 405.
[64] P. Anooj, Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules and decision tree rules, Open Comput. Sci. 1 (2011) 482–498.
[65] R. Arsanjani, Y. Xu, D. Dey, V. Vahistha, A. Shalev, R. Nakanishi, S. Hayes, M. Fish, D. Berman, G. Germano, P.J. Slomka, Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population, J. Nucl. Cardiol. 20 (2013) 553–562.
[66] L. Verma, S. Srivastava, P.C. Negi, A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data, J. Med. Syst. 40 (2016) 1–7.
[67] Y. Yang, L. Chen, Y. Yam, S. Achenbach, M. Al-Mallah, D.S. Berman, M.J. Budoff, F. Cademartiri, T.Q. Callister, H.-J. Chang, V.Y. Cheng, K. Chinnaiyan, R. Cury, A. Delago, A. Dunning, G. Feuchtner, M. Hadamitzky, J. Hausleiter, R.P. Karlsberg, P.A. Kaufmann, Y.-J. Kim, J. Leipsic, T. LaBounty, F. Lin, E. Maffei, G.L. Raff, L.J. Shaw, T.C. Villines, J.K. Min, B.J.W. Chow, A clinical model to identify patients with high-risk coronary artery disease, JACC (J. Am. Coll. Cardiol.): Cardiovasc. Imaging 8 (2015) 1–9.
[68] I. Mandal, N. Sairam, Accurate prediction of coronary artery disease using reliable diagnosis system, J. Med. Syst. 36 (2012) 3353–3373.
[69] A. Lahsasna, R.N. Ainon, R. Zainuddin, A. Bulgiba, Design of a fuzzy-based decision support system for coronary heart disease diagnosis, J. Med. Syst. 36 (2012) 3293–3306.
[70] A.Q. Ansari, N.K. Gupta, Automated Diagnosis of Coronary Heart Disease Using Neuro-Fuzzy Integrated System, World Congress on Information and Communication Technologies, 2011, pp. 1379–1384.
[71] E.I. Georga, N.S. Tachos, A.I. Sakellarios, V.I. Kigka, T.P. Exarchos, G. Pelosi, O. Parodi, L.K. Michalis, D.I. Fotiadis, Artificial Intelligence and Data Mining Methods for Cardiovascular Risk Prediction, Cardiovascular

Computing—Methodologies and Clinical Applications, Springer2019, pp. 279-301.
[72] E.M. Karabulut, T. İbrikçi, Effective diagnosis of coronary artery disease using the rotation forest ensemble method, J. Med. Syst. 36 (2012) 3011–3018.
[73] B.L.D. Akhil jabbar, M. Priti Chandra, Classification of heart disease using artificial neural network and feature subset selection, Glob. J. Comput. Sci. Technol. 8 (2013) 5–14.
[74] H. Uğuz, A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases, J. Med. Syst. 36 (2012) 61–72.
[75] H.G. Lee, W.-S. Kim, K.Y. Noh, J.-H. Shin, U. Yun, K.H. Ryu, Coronary artery disease prediction method using linear and nonlinear feature of heart rate variability in three recumbent postures, Inf. Syst. Front. 11 (2009) 419–431.
[76] J.-K. Kim, J.-S. Lee, D.-K. Park, Y.-S. Lim, Y.-H. Lee, E.-Y. Jung, Adaptive mining prediction model for content recommendation to coronary heart disease patients, Clust. Comput. 17 (2014) 881–891.
[77] J.-G. Yang, J.-K. Kim, U.-G. Kang, Y.-H. Lee, Coronary heart disease optimization system on adaptive-network-based fuzzy inference system and linear discriminant analysis (ANFIS—LDA), Personal Ubiquitous Comput. 18 (2014) 1351–1362.
[78] S. Fujimoto, T. Kondo, H. Yamamoto, N. Yokoyama, Y. Tarutani, K. Takamura, Y. Urabe, K. Konno, Y. Nishizaki, T. Shinozaki, Y. Kihara, H. Daida, T. Isshiki, S. Takase, Development of new risk score for pre-test probability of obstructive coronary artery disease based on coronary CT angiography, Heart Vessel. 30 (2015) 563–571.
[79] G. Parthiban, A. Rajesh, S. Srivatsa, Diagnosis of heart disease for diabetic patients using naive bayes method, Int. J. Comput. Appl. 24 (2011) 7–11.
[80] G. Parthiban, S. Srivatsa, Applying machine learning methods in diagnosing heart disease for diabetic patients, Int. J. Appl. Inf. Sys. (IJAIS) 3 (2012) 2249-0868.
[81] A.H. Chen, S.Y. Huang, P.S. Hong, C.H. Cheng, E.J. Lin, HDPS: heart disease prediction system, Comput. Cardiol. (2011) 557–560.
[82] S. Muthukaruppan, M.J. Er, A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease, Expert Syst. Appl. 39 (2012) 11657–11665.
[83] G.S.R. Rajkumar, Asha, Diagonsis of heaer disease using datamining algorithm, Glob. J. Comput. Sci. Technol. 10 (2010) 38–43.
[84] U.R. Acharya, H. Fujita, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, R.S. Tan, Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals, Appl. Intell. 49 (2019) 16–27.
[85] U.R. Acharya, Y. Hagiwara, J.E.W. Koh, S.L. Oh, J.H. Tan, M. Adam, R.S. Tan, Entropies for automated detection of coronary artery disease using ECG signals, Biocybern. Biomed. Eng. 38 (2018) 373–384.
[86] G.-P. Liu, G.-Z. Li, Y.-L. Wang, Y.-Q. Wang, Modelling of inquiry diagnosis for coronary heart disease in traditional Chinese medicine by using multi-label learning, BMC Complement Altern. Med. 10 (2010) 1–12.
[87] E. Ahmadi, G.R. Weckman, D.T. Masel, Decision making model to predict presence of coronary artery disease using neural network and C5.0 decision tree, J. Ambient Intell. Humanized Comput. 9 (2018) 999–1011.
[88] M. Shouman, T. Turner, R. Stocker, Integrating Naive Bayes and K-Means Clustering with Different Initial Centroid Selection Methods in the Diagnosis of Heart Disease Patients, CS & IT-CSCP, 2012, pp. 125–137.
[89] C. Caselli, D. Rovai, V. Lorenzoni, C. Carpeggiani, A. Teresinska, S. Aguade, G. Todiere, A. Gimelli, S. Schroeder, G. Casolo, R. Poddighe, F. Pugliese, D. Le Guludec, S. Valente, G. Sambuceti, P. Perrone-Filardi, S. Del Ry, M. Marinelli, S. Nekolla, M. Pietila, M. Lombardi, R. Sicari, A. Scholte, J. Zamorano, P.A. Kaufmann, S.R. Underwood, J. Knuuti, D. Giannessi, D. Neglia, A new integrated clinical-biohumoral model to predict functionally significant coronary artery disease in patients with chronic chest pain, Can. J. Cardiol. 31 (2015) 709–716.
[90] L. Verma, S. Srivastava, A data mining model for coronary artery disease detection using noninvasive clinical parameters, Indian J. Sci. Technol. 9 (2016) 1–6.
[91] M.A. Jabbar, B. Deekshatulu, P. Chandra, Heart disease classification using nearest neighbor classifier with feature subset selection, Anale. Seria Informatica 11 (2013) 47–54.
[92] M. Karaolis, J.A. Moutiris, C.S. Pattichis, Assessment of the risk of coronary heart event based on data mining, 8th IEEE International Conference on BioInformatics and BioEngineering, 2008, pp. 1–5.
[93] L. Verma, S. Srivastava, P.C. Negi, An intelligent noninvasive model for coronary artery disease detection, Complex. Intell. Syst. 4 (2018) 11–18.
[94] U.R. Acharya, H. Fujita, O.S. Lih, M. Adam, J.H. Tan, C.K. Chua, Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network, Knowl. Based Syst. 132 (2017) 62–71.
[95] U.R. Acharya, H. Fujita, M. Adam, O.S. Lih, V.K. Sudarshan, T.J. Hong, J.E.W. Koh, Y. Hagiwara, C.K. Chua, C.K. Poo, T.R. San, Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of ECG signals: a comparative study, Inf. Sci. 377 (2017) 17–29.
[96] R. Alizadehsani, M.H. Zangooei, M.J. Hosseini, J. Habibi, A. Khosravi, M. Roshanzamir, F. Khozeimeh, N. Sarrafzadegan, S. Nahavandi, Coronary artery disease detection using computational intelligence methods, Knowl. Based Syst. 109 (2016) 187–197.
[97] I. Babaoglu, O.K. Baykan, N. Aygul, K. Ozdemir, M. Bayrak, Assessment of exercise stress testing with artificial neural network in determining coronary artery disease and predicting lesion localization, Expert Syst. Appl. 36 (2009) 2562–2566.
[98] K. Buchan, M. Filannino, Ö. Uzuner, Automatic prediction of coronary artery disease from clinical narratives, J. Biomed. Inform. 72 (2017) 23–32.
[99] X. Liu, H. Fu, PSO-based support vector machine with cuckoo search technique for clinical disease diagnoses, Sci. World J. 2 (2014) 1–7.
[100] J.K. Kim, S. Kang, Neural network-based coronary heart disease risk prediction using feature correlation analysis, J. Healthc. Eng. (2017) 1–13.
[101] W.M. Baihaqi, N.A. Setiawan, I. Ardiyanto, Rule Extraction for Fuzzy Expert System to Diagnose Coronary Artery Disease, 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2016, pp. 136–141.
[102] M. Akay, Noninvasive diagnosis of coronary artery disease using a neural network algorithm, Biol. Cybern. 67 (1992) 361–367.
[103] I. Kurt, M. Ture, A.T. Kurum, Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease, Expert Syst. Appl. 34 (2008) 366–374.
[104] R. Chitra, V. Seenivasagam, Heart disease prediction system using supervised learning classifier, Bonfring Int. J. Comput. Eng. Soft Comput. 3 (2013) 01–07.
[105] U. Shafique, F. Majeed, H. Qaiser, I.U. Mustafa, Data mining in healthcare for heart diseases, Int. J. Innov. Appl. Stud. 10 (2015) 1312–1322.
[106] K. Polat, S. Şahan, S. Güneş, Automatic detection of heart disease using an artificial immune recognition system (AIRS) with fuzzy resource allocation mechanism and k-nn (nearest neighbour) based weighting preprocessing, Expert Syst. Appl. 32 (2007) 625–631.
[107] D. Tomar, S. Agarwal, Feature selection based least square twin support vector machine for diagnosis of heart disease, Int. J. Bio Sci. Bio Technol. 6 (2014) 69–82.
[108] U.R. Acharya, H. Fujita, V.K. Sudarshan, S.L. Oh, M. Adam, J.H. Tan, J.H. Koo, A. Jain, C.M. Lim, K.C. Chua, Automated characterization of coronary artery disease, myocardial infarction, and congestive heart failure using contourlet and shearlet transforms of electrocardiogram signal, Knowl. Based Syst. 132 (2017) 156–166.
[109] Y.N. Devi, S. Anto, An evolutionary-fuzzy expert system for the diagnosis of coronary artery disease, Int. J. Adv. Res. Comput. Eng. Technol. 3 (2014) 2278-1323.
[110] M. Kumar, R.B. Pachori, U.R. Acharya, Characterization of coronary artery disease using flexible analytic wavelet transform applied on ECG signals, Biomed. Signal Process. Control 31 (2017) 301–308.
[111] L. Huan, R. Setiono, Feature selection via discretization, IEEE Trans. Knowl. Data Eng. 9 (1997) 642–645.
[112] U.R. Acharya, V.K. Sudarshan, J.E.W. Koh, R.J. Martis, J.H. Tan, S.L. Oh, A. Muhammad, Y. Hagiwara, M.R.K. Mookiah, K.P. Chua, C.K. Chua, R.S. Tan, Application of higher-order spectra for the characterization of Coronary artery disease using electrocardiogram signals, Biomed. Signal Process. Control 31 (2017) 31–43.
[113] R. Alizadehsani, H. Mohammad Javad, B. Reihane, G. Asma, K. Fahime, S. Zahra Alizadeh, Exerting cost-sensitive and feature creation algorithms for coronary artery disease diagnosis, Int. J. Knowl. Discov. Bioinform. 3 (2012) 59–79.
[114] A. Dekamin, A. Sheibatolhamdi, A data mining approach for coronary artery disease prediction in Iran, J. Adv. Med. Sci. Appl. Technol. 3 (2017) 29–38.
[115] A.D. Souza, Heart disease prediction using data mining techniques, Int. J. Res. Eng. Sci 3 (2015) 74–77.
[116] M. Anbarasi, E. Anupriya, N. Iyengar, Enhanced prediction of heart disease with feature subset selection using genetic algorithm, Int. J. Eng. Sci. Technol. 2 (2010) 5370–5376.
[117] K. Vembandasamy, R. Sasipriya, E. Deepa, Heart diseases detection using naive bayes algorithm, IJISET Int. J. Innov. Sci. Eng. Technol. 2 (2015) 441–444.
[118] D. Giri, U. Rajendra Acharya, R.J. Martis, S. Vinitha Sree, T.-C. Lim, T. Ahamed, J.S. Suri, Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform, Knowl. Based Syst. 37 (2013) 274–282.
[119] R. Khemchandani, S. Chandra Jayadeva, Fast and robust learning through fuzzy linear proximal support vector machines, Neurocomputing 61 (2004) 401–411.
[120] C. Ordonez, Comparing association rules and decision trees for disease prediction, Proceedings of the international workshop on Healthcare information and knowledge management, 2006, pp. 17–24.
[121] H. Yang, I. Spasic, J.A. Keane, G. Nenadic, A text mining approach to the prediction of disease status from clinical discharge summaries, J. Am. Med. Inform. Assoc. 16 (2009) 596–600.
[122] M. Shouman, T. Turner, R. Stocker, Applying k-nearest neighbour in diagnosing heart disease patients, Int. J. Inf. Educ. Technol. 2 (2012) 220–223.
[123] S. Sood, M. Kumar, R.B. Pachori, U.R. Acharya, Application of empirical mode decomposition–based features for analysis of normal and cad heart rate signals, J. Mech. Med. Biol. 16 (2016) 1–20.
[124] V. Chaurasia, S. Pal, Early prediction of heart diseases using data mining techniques, Caribb. J. Sci. Technol. 1 (2013) 208–217.
[125] M. Nalluri, D. Roy, Hybrid disease diagnosis using multiobjective optimization with evolutionary parameter optimization, J. Healthc. Eng. 2 (2017) 1–27.
[126] L. Tapak, H. Mahjub, O. Hamidi, J. Poorolajal, Real-data comparison of data mining methods in prediction of diabetes in Iran, Healthc. Inf. Res. 19 (2013) 177–185.
[127] T. Nadu, Cad diagnosis using PSO, BAT, MLR and SVM, Int. J. Adv. Res. Ideas Innov. Technol. 3 (2017) 690–694.
[128] D. Prabowo, N. Setiawan, H. Nugroho, A study of data randomization on a computer based feature selection for diagnosing coronary artery disease, Adv. Intell. Syst. 53 (2014) 237–248.
[129] S. Palaniappan, R. Awang, Intelligent heart disease prediction system using data mining techniques, IEEE/ACS Int. Conf. Comput. Syst. Appl. 8 (2008) 108–115.
[130] N. Ikeda, K. Sugi, M. Nakamura, U.R. Acharya, L. Saba, G. Mercuro, F. Molinari, A. Nicolaides, J.S. Suri, Abstract 13515: a feature classification approach for coronary artery disease prediction via carotid atherosclerosis window, Circulation 128 (2013) A13515.
[131] M. Abdar, Using decision trees in data mining for predicting factors influencing of

heart disease, Carpathian J. Electron. Comput. Eng. 8 (2015) 31–36.
[132] C.K. Tham, C.K. Heng, W.C. Chin, Predicting risk of coronary artery disease from DNA microarray-based genotyping using neural networks and other statistical analysis tool, J. Bioinform. Comput. Biol. 01 (2003) 521–539.
[133] R. Alizadehsani, M.J. Hosseini, Z. Alizadehsani, M.H. Mohammadi, O. Barati, F. Khozeimeh, System for Determining the Need for Angiography in Patients with Symptoms of Coronary Artery Disease, US Patents, 2015, pp. 1–8.
[134] U. Raghavendra, H. Fujita, A. Gudigar, R. Shetty, K. Nayak, U. Pai, J. Samanth, U.R. Acharya, Automated technique for coronary artery disease characterization and classification using DD-DTDWT in ultrasound images, Biomed. Signal Process. Control 40 (2018) 324–334.
[135] A. Patel, D. Shukla, A data mining technique for prediction of coronary heart disease using neuro-fuzzy integrated approach two level, Int. J. Eng. Math. Sci. 2 (2013) 2663–2671.
[136] O.Y. Atkov, S.G. Gorokhova, A.G. Sboev, E.V. Generozov, E.V. Muraseyeva, S.Y. Moroshkina, N.N. Cherniy, Coronary heart disease diagnosis by artificial neural networks including genetic polymorphisms and clinical parameters, J. Cardiol. 59 (2012) 190–194.
[137] J.H. Tan, Y. Hagiwara, W. Pang, I. Lim, S.L. Oh, M. Adam, R.S. Tan, M. Chen, U.R. Acharya, Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals, Comput. Biol. Med. 94 (2018) 19–26.
[138] S. Patidar, R.B. Pachori, U. Rajendra Acharya, Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals, Knowl. Based Syst. 82 (2015) 1–10.
[139] S. Dua, X. Du, Novel classification of coronary artery disease using heart rate variability analysis, J. Mech. Med. Biol. 12 (2012) 1–19.
[140] J. Nair, M. Ghatge, V. Kakkar, J. Shanker, Network analysis of inflammatory genes and their transcriptional regulators in coronary artery disease, PLoS One 9 (2014) e94328.
[141] E.V. Garcia, C.D. Cooke, R.D. Folks, C.A. Santana, E.G. Krawczynska, L. De Braal, N.F. Ezquerra, Diagnostic performance of an expert system for the interpretation of myocardial perfusion SPECT studies, J. Nucl. Med. 42 (2001) 1185–1191.
[142] H. Moghaddasi, B. Ahmadzadeh, R. Rabiei, M. Farahbakhsh, Study on the efficiency of a multi-layer perceptron neural network based on the number of hidden layers and nodes for diagnosing coronary-artery disease, Jentashapir J. Health Res. 8 (2017) 1–8.
[143] L. Kaur, Predicting heart disease symptoms using fuzzy C-means clustering, Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) 3 (2014) 4232–4235.
[144] R.A. Mohammadpour, S.M. Abedi, S. Bagheri, A. Ghaemian, Fuzzy rule-based classification system for assessing coronary artery disease, Comput. Math. Math. Med. 3 (2015) 1–8.
[145] R. Rajkumar, K. Ananadakumar, A. Bharathi, Risk level classification of coronary artery heart disease in diabetic patients using neuro fuzzy classifier, Int. J. Comput. Intell. Res. 13 (2017) 575–582.
[146] J.G. Bazan, S. Buregwa-Czuma, P.W. Pardel, S. Bazan-Socha, B. Sokołowska, S. Dziedzina, Predicting the presence of serious coronary artery disease based on 24 hour holter ECG monitoring, Trans. Rough Sets (2015) 95–113.
[147] W. Wiharto, H. Kusnanto, H. Herianto, Interpretation of clinical data based on C4.5 algorithm for the diagnosis of coronary heart disease, Healthc. Inform. Res. 22 (2016) 186–195.
[148] M. Kumar, R.B. Pachori, U. Rajendra Acharya, An efficient automated technique for CAD diagnosis using flexible analytic wavelet transform and entropy features extracted from HRV signals, Expert Syst. Appl. 63 (2016) 165–172.
[149] R. Alizadehsani, J. Habibi, Z.A. Sani, H. Mashayekhi, R. Boghrati, A. Ghandeharioun, B. Bahadorian, Diagnosis of coronary artery disease using data mining based on lab data and echo features, J. Med. Bioeng. 1 (2012) 26–29.
[150] C.-M. Chu, W.-C. Chien, C.-H. Lai, H.B. Bludau, H.-J. Tschai, L. Pai, S.-M. Hsieh, N.F. Chu, A. Klar, R. Haux, A Bayesian expert system for clinical detecting coronary artery disease, J. Med. Sci. 29 (2009) 187–194.
[151] M.G. Tsipouras, T.P. Exarchos, D.I. Fotiadis, A. Kotsia, A. Naka, L.K. Michalis, A decision support system for the diagnosis of coronary artery disease, 19th IEEE Symposium on Computer-Based Medical Systems, 2006, pp. 279–284 CBMS'06.
[152] R. Alizadehsani, J. Habibi, M.J. Hosseini, R. Boghrati, A. Ghandeharioun, B. Bahadorian, Z.A. Sani, Diagnosis of coronary artery disease using data mining techniques based on symptoms and ecg features, Eur. J. Sci. Res. 82 (2012) 542–553.
[153] C. Yadav, S. Lade, M.K. Suman, Predictive analysis for the diagnosis of coronary artery disease using association rule mining, Int. J. Comput. Appl. 87 (2014) 9–13.
[154] R. Das, I. Turkoglu, A. Sengur, Effective diagnosis of heart disease through neural networks ensembles, Expert Syst. Appl. 36 (2009) 7675–7680.
[155] A.K. Paul, P.C. Shill, M.R.I. Rabin, K. Murase, Adaptive weighted fuzzy rule-based system for the risk level assessment of heart disease, Appl. Intell. 48 (2018) 1739–1756.
[156] S. Iftikhar, K. Fatima, A. Rehman, A.S. Almazyad, T. Saba, An evolution based hybrid approach for heart diseases classification and associated risk factors identification, Biomed. Res. 28 (2017) 3451–3455.
[157] N. Setiawan, P. Venkatachalam, A.M. Fadzil, Rule selection for coronary artery disease diagnosis based on rough set, Int. J. Recent Trends Eng. 2 (2009) 198–202.
[158] N. Süt, M. Şenocak, Assessment of the performances of multilayer perceptron neural networks in comparison with recurrent neural networks and two statistical methods for diagnosing coronary artery disease, Expert Syst. 24 (2007) 131–142.
[159] M. Tayefi, M. Tajfard, S. Saffar, P. Hanachi, A.R. Amirabadizadeh, H. Esmaeily, A. Taghipour, G.A. Ferns, M. Moohebati, M. Ghayour-Mobarhan, hs-CRP is strongly associated with coronary heart disease (CHD): a data mining approach using decision tree algorithm, Comput. Methods Progr. Biomed. 141 (2017) 105–109.
[160] C. Sridhar, U.R. Acharya, H. Fujita, G.M. Bairy, Automated diagnosis of Coronary Artery Disease using nonlinear features extracted from ECG signals, IEEE Int. Conf. Syst. Man Cybern. (2016) 545–549.
[161] R. Alizadehsani, M.J. Hosseini, Z.A. Sani, A. Ghandeharioun, R. Boghrati, Diagnosis of coronary artery disease using cost-sensitive algorithms, IEEE 12th International Conference on Data Mining Workshops, 2012, pp. 9–16.
[162] S. Nithya, Bat Imperialist Competitive Algorithm (BICA) Based Feature Selection and Genetic Fuzzy Based Improved Kernel Support Vector Machine (GF-IKSVM) Classifier for Diagnosis of Cardiovascular Heart Disease, Biomedical Research, 2018, pp. S95–S104.
[163] N. Allahverdi, G. Altan, Y. Kutlu, Diagnosis of coronary artery disease using deep belief networks, 2, Int. Conf. Eng. Nat. Sci (2016) 1–9.
[164] M. Abdar, S.R.N. Kalhori, T. Sutikno, I.M.I. Subroto, G. Arji, Comparing performance of data mining algorithms in prediction heart diseases, Int. J. Electr. Comput. Eng. 5 (2015) 1569–1576.
[165] M. Negahbani, S. Joulazadeh, H. Marateb, M. Mansourian, Coronary artery disease diagnosis using supervised fuzzy c-means with differential search algorithm-based generalized Minkowski metrics, Peertechz J. Biomed. Eng. 1 (1) (2015) 006, 14.
[166] H. Moghaddasi, I. Mahmoudi, S. Sajadi, Comparing the efficiency of artificial neural network and gene expression programming in predicting coronary artery disease, J. Health Med. Inform. 8 (2017) 1–6.
[167] R. Alizadehsani, J. Habibi, Z. Alizadeh Sani, H. Mashayekhi, R. Boghrati, A. Ghandeharioun, F. Khozeimeh, F. Alizadeh-Sani, Diagnosing coronary artery disease via data mining algorithms by considering laboratory and echocardiography features, Res. Cardiovasc. Med. 2 (2013) 133–139.
[168] R. Alizadehsani, J. Habibi, B. Bahadorian, H. Mashayekhi, A. Ghandeharioun, R. Boghrati, Z.A. Sani, Diagnosis of coronary arteries stenosis using data mining, J. Med. Signals Sens. 2 (2012) 153–159.
[169] J. Chen, P. Lu, X. Zuo, Q. Shi, H. Zhao, L. Luo, J. Yi, C. Zheng, Y. Yang, W. Wang, Clinical data mining of phenotypic network in angina pectoris of coronary heart disease, Evid. Based Complement Altern. Med. 2 (2012) 1–8.
[170] W. Liu, X. Xiong, X. Yang, F. Chu, H. Liu, The effect of Chinese herbal medicine gualouxiebaibanxia decoction for the treatment of angina pectoris: a systematic review, Evid. Based Complement Altern. Med. 4 (2016) 1–14.
[171] Y. Zhao, X. Yu, X. Cao, L. Luo, L. He, S. Mao, L. Ma, P. Rong, Y. Zhao, G. Li, B. Liu, Cluster analysis for syndromes of real-world coronary heart disease with angina pectoris, Front. Med. (2017) 1–6.
[172] U.R. Acharya, S.V. Sree, M. Muthu Rama Krishnan, N. Krishnananda, S. Ranjan, P. Umesh, J.S. Suri, Automated classification of patients with coronary artery disease using grayscale features from left ventricle echocardiographic images, Comput. Methods Progr. Biomed. 112 (2013) 624–632.
[173] U.R. Acharya, O. Faust, V. Sree, G. Swapna, R.J. Martis, N.A. Kadri, J.S. Suri, Linear and nonlinear analysis of normal and CAD-affected heart rate signals, Comput. Methods Progr. Biomed. 113 (2014) 55–68.
[174] M. Sharma, U. Rajendra Acharya, A New Method to Identify Coronary Artery Disease with ECG Signals and Time-Frequency Concentrated Antisymmetric Biorthogonal Wavelet Filter Bank, Pattern Recognition Letters, 2019.
[175] I. Chadès, S. Nicol, Small data call for big ideas, Nature 539 (2016) 1–5.
[176] T. Wang, Z. Qin, Z. Jin, S. Zhang, Handling over-fitting in test cost-sensitive decision tree learning by feature selection, smoothing and pruning, J. Syst. Softw. 83 (2010) 1137–1147.
[177] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3 (2003) 1157–1182.
[178] M. Dash, H. Liu, Feature selection for classification, Intell. Data Anal. 1 (1997) 131–156.
[179] T. Mitchell, B. Buchanan, G. DeJong, T. Dietterich, P. Rosenbloom, A. Waibel, Machine learning, Annu. Rev. Comput. Sci. 4 (1990) 417–433.
[180] Y. Shen, Z. Fang, Y. Gao, N. Xiong, C. Zhong, X. Tang, Coronary arteries segmentation based on 3D FCN with attention gate and level set function, IEEE Access 7 (2019) 42826–42835.
[181] U.R. Acharya, H. Fujita, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals, Inf. Sci. 415–416 (2017) 190–198.
[182] J. Betancur, F. Commandeur, M. Motlagh, T. Sharir, A.J. Einstein, S. Bokhari, M.B. Fish, T.D. Ruddy, P. Kaufmann, A.J. Sinusas, Deep learning for prediction of obstructive disease from fast myocardial perfusion, SPECT, JACC: Cardiovasc. Imaging 11 (2018) 1654–1664.
[183] J. Betancur, L.-H. Hu, F. Commandeur, T. Sharir, A.J. Einstein, M.B. Fish, Deep learning analysis of upright-supine high-efficiency spect myocardial perfusion imaging for prediction of obstructive coronary artery disease: a multicenter study, J. Nucl. Med. 60 (2019) 664–670.
[184] A. Abdolmanafi, L. Duong, N. Dahdah, I.R. Adib, F. Cheriet, Characterization of coronary artery pathological formations from OCT imaging using deep learning, Biomed. Opt. Express 9 (2018) 4936–4960.
[185] A. Yeri, R.V. Shah, Comparison of computational fluid dynamics and machine learning based fractional flow reserve in coronary artery disease, circulation, Cardiovasc. Imaging 11 (2018) e007950.
[186] J. Rubin, R. Abreu, A. Ganguli, S. Nelaturi, I. Matei, K. Sricharan, Recognizing Abnormal Heart Sounds Using Deep Learning, (2017), pp. 1–7 arXiv preprint arXiv:.04642.
[187] R.W. van Hamersvelt, M. Zreik, M. Voskuil, M.A. Viergever, I. Išgum, T. Leiner, Deep learning analysis of left ventricular myocardium in CT angiographic intermediate-degree coronary stenosis improves the diagnostic accuracy for identification of functionally significant stenosis, Eur. Radiol. 29 (2019) 2350–2359.
[188] H. Sofian, J.T.C. Ming, S. Mohamad, N.M. Noor, Calcification detection using deep

structured learning in intravascular ultrasound image for coronary artery disease, 2nd International Conference on BioSignal Analysis, Processing and Systems, ICBAPS, 2018, pp. 47–52.
[189] M. Zreik, N. Lessmann, R.W. van Hamersvelt, J.M. Wolterink, M. Voskuil, M.A. Viergever, T. Leiner, I. Išgum, Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis, Med. Image Anal. 44 (2018) 72–85.
[190] M.A. Mazurowski, P.A. Habas, J.M. Zurada, J.Y. Lo, J.A. Baker, G.D. Tourassi, Training neural network classifiers for medical decision making: the effects of imbalanced datasets on classification performance, Neural Network. 21 (2008) 427–436.
[191] N. Arsov, M. Pavlovski, L. Basnarkov, L. Kocarev, Generating highly accurate prediction hypotheses through collaborative ensemble learning, Sci. Rep. 7 (2017) 1–9.
[192] D.H. Wolpert, W.G. Macready, No free lunch theorems for optimization, IEEE Trans. Evol. Comput. 1 (1997) 67–82.
[193] J.V. Tu, Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes, J. Clin. Epidemiol. 49 (1996) 1225–1231.
[194] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (2017) 115–118.
[195] I. Inza, P. Larrañaga, Y. Saeys, A review of feature selection techniques in bioinformatics, Bioinformatics 23 (2007) 2507–2517.
[196] T. Bäck, D.B. Fogel, Z. Michalewicz, Evolutionary Computation 1: Basic Algorithms and Operators, CRC press, 2018.
[197] O. Faust, Y. Hagiwara, T.J. Hong, O.S. Lih, U.R. Acharya, Deep learning for healthcare applications based on physiological signals: a review, Comput. Methods Progr. Biomed. 161 (2018) 1–13.