A SYSTEMATIC LITERATURE REVIEW OF DEEP AND MACHINE LEARNING ALGORITHMS IN CARDIOVASCULAR DISEASES DIAGNOSIS

ZAKARIA K. D. ALKAYYALI [1], SYAHRIL ANUAR BIN IDRIS [2], SAMY S. ABU-NASER [3]

[1,2] University Malaysia of Computer Science & Engineering (UNIMY), Cyberjaya, Malaysia

[3]Faculty of Engineering and Information Technology, Al-Azhar University, Gaza, Palestine

Email: P02220001@student.unimy.edu.my[1], zkayyali@qou.edu[1], syahril.idris@unimy.edu.my[2], abunaser@alazhar.edu.ps[3]

ABSTRACT

During the whole cardiac cycle, heart sounds are created, and blood enters the heart chambers as the cardiac regulators open and close. Blood flow produces aural noises; the more turbulent the blood flow, the more ambiances are created. Two common cardiac sounds occur in sequence with each heartbeat in healthy adults. These are the first heart sound (S1) and second heart sound (S2), which are caused by the closure of the atrioventricular and semilunar valves, respectively. The current systematic review depends on "the Preferred Reporting Items for Systematic reviews and Meta-Analysis statement" and 40 appropriate studies. The search of the literature employed search engines similar to: IEEE Xplore, Google Scholar, Hindawi, PubMed, SCOPUS, Wiley Online, Web of Science, Taylor and Francis, ScienceDirect, and Ebscohost. This study concentrated on four characteristics: Algorithms of Machine and Deep Learning, best-algorithm performance, datasets, and application used in Cardiovascular Diseases predictions. The experimental articles did not use Reinforcement Learning, Semi-supervised learning, promising aspects of Deep and Machine Learning. Algorithms based on ensemble technique exhibited sensible rates of accuracy nonetheless were not frequent, whereas Convolutional Neural Network (CNN) were well epitomized. A few studies smeared main datasets (13 of 37). Recurrent Neural Network (RNN), boosting algorithms, Support Vector Machine (SVM) and K-Nearest Neighbors (KNN), were the best performing algorithms. This review will be beneficial for investigators predicting Cardiovascular Diseases using machine and deep learning methods.

Keywords—*Cardiovascular Diseases, Datasets, Algorithms, Deep Learning, Machine Learning*

1. INTRODUCTION

Globally, chronic diseases including diabetes, heart disease, and cancer are the main killers. The healthcare industry is completely unique from all other industries. Consumers, regardless of price, demand the best level of care and services in this high priority area. The leading cause of mortality is cardiovascular disease (CVDs). More individuals worldwide die each year from CVDs than any other cause. According to estimates, coronary heart disorders caused 17.9 million CVD deaths worldwide in 2019. During the cardiac cycle, heart sounds are produced as blood enters the heart chambers as the cardiac valves open and close.

Auditory noises are produced by the blood flow, and the more turbulent the blood flow, the more vibrations are generated There are two typical heart sounds that occur with each heartbeat in healthy adults.

Prompt discovery and diagnosis are acute for decreasing the Cardiovascular Diseases. Artificial-intelligence (AI), a subfield of computer science, places a strong emphasis on the creation of hardware and software with intelligence that resembles human behavior. Computer programs that have been upgraded with artificial intelligence are capable of learning, planning, and problem-solving activities, among others.

Benefits of AI are numerous as recognized in the literature. These comprise helping specialists do complex operations, appropriate decision-making, and jobs, to provide precise Cardiovascular Diseases in images, to decrease the threats of composite handlings, to enhance Cardiovascular Diseases knowledge about individual behavior, and enhancing computer assistance identification [3]. AI, Machine Learning (ML) and Deep Learning (DL) research concentrate on in-expensive, fast, and none invasive approaches to precisely detect Cardiovascular Diseases using advanced metrics of performance likes: sensitivity, Recall, Accuracy, F1-Score, Precision, and specificity [4].

Machine and deep learning permit PCs to discover, calculate, and understand relationships among attributes to increase preventive medicine by analytically learning ideal data representations [5]. Deep learning processes can examine through substantial volumes of Cardiovascular Diseases data, permitting it to discover prognostic, diagnostic, and remedial treatment options for various Cardiovascular Diseases easily. The most popular types of deep learning techniques are: supervised, un-supervised, reinforcement and Semi-supervised learnings [6]. Supervised-learning is a vigorous method that uses computer language to categorize and understand labeled Cardiovascular Diseases data [7]. Such as, in supervised learning, a specialist may pursue to identify whether an image represents benign, malignant or pituitary? Consequently, supervised-learning needs a dataset with images and pre-defined labels [8]. Unsupervised-learning aims to discover the principal construction or relationships among attributes in a given data set [9]. This data set will be trained with no labels for the images, and the model gathers the data to categorize the essential arrangements. According to behavioral consciousness, reinforcement-learning uses another approach in which a software functions in a pre-determined setting to maximize a return. Semi-supervised-learning is a deep learning method that blends a large volume of unlabeled data with a small volume of labeled data throughout the training. It lies between un-supervised and supervised learning.

The focal aim of the current Systematic Literature Review (SLR) is to identify which supervised-learning procedures show the top outcomes for Cardiovascular Diseases forecasting. In current SLR: 1) the researchers identified articles that used deep-learning methods to identify Cardiovascular Diseases 2) to recognize the top employed supervised-deep-learning procedures for CVD forecasting; 3) to assess the supervised deep learning procedures performance of in relation to the designated measures for example sensitivity, F1-Score, Recall, specificity, precision, and Accuracy; 4) to investigate the data sets for the prediction of Cardiovascular Diseases. The results of this SLR will give authors the guideline, training, and additional research on Cardiovascular Diseases. The reset of the paper is organized like this: Section 2 will present the Methods and Materials, Section 3, outline the Review of the Literature, Section 4, provide detailed Discussions, and the concluding section present the Conclusion and futures of work [10].

2. METHODS AND MATERIALS

The embraced procedure in this SLR is "the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) statement". The PRISMA is a group of components employed to describe SLR and analyses [11]. It is envisioned to support report analyses and arbitraries trial measurement [12], furthermore, it could be used as a basis for a systematic reviews reporting [13]. By applying the PRISMA method, this current study took into consideration a few research questions directing the SLR, criteria for the searching of literature, and selection criteria.

2.1 Literature Search Criteria

The criteria of selection recognized important review articles and published research by means of key terms for example Cardiovascular, Diseases, Cardiovascular Diseases, Cardiovascular problems, artificial intelligence, deep learning algorithms, deep learning techniques, data mining, machine learning techniques, machine learning methods, and in Cardiovascular Diseases. The study embraced combined search criteria literature that associates key terms by applying Boolean operators of Boolean like "and, or, --, ~". Among the search engines employed includes: ScienceDirect, Ebscohost, Google Scholar, SCOPUS, Hindawi, Web of Science, Wiley Online, IEEE, PubMed, Taylor and Francis. Thus, the search result produced approximately 520 research articles, 37 of them were considered suitable depending on the selection criteria in the current study.

2.2 Selection Criteria

Our study employed published research papers written in the English language. Table 1 shows inclusion and exclusion criterion employed in our study. The found research articles concentrated on the Deep and Machine Learning application methods in examining Cardiovascular Diseases. The criteria selection omitted non-research papers like Thesis, chapters, books, and were omitted from the SLR.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Paper selection criteria

Feature	Criteria Inclusion	Criteria Exclusion
Publication Language	English	Not English
Type of Research	Research Papers	Thesis/Dissertation, case studies, books, reports, and magazines.
Focus of Research	Techniques related to DL, ML in Cardiovascular Diseases	Not in relation with DL, ML and CVD
Setting	Worldwide	(N/A)

3. RESULTS

A search of the literature in aforementioned databases recognized 520 articles (as in Figure 1). Sorting via abstract indicates 240 articles did not match the criteria selection for the current literature review. Further sorting by title showed 160 articles did not comply with the focus of the systematic review. On the other hand, complete text sorting of the leftover articles show 45 research papers were connected and they were not using DL and ML methods to CVD. Consequently, those articles were left out from the last listing. To conclude, 37 research papers on the application of DL and ML procedures in CVD were contained within the current review.

The list of references in Table 2 shows 37 articles used in the current SLR. From the previous studies, the focus area of those papers were: 1) Prediction of Heart disease detection, 2) detection of Heart disease, and 3) diagnosis of Heart disease by DL and ML methods. A substantial quantity of those articles meant at heart sound prediction methods. The collected papers covered the years from 2018-2022. Figure 2 portrays the ratio of the gathered research papers on the occurrence over the time. The count of research papers has varied gradually with time with the exception of from 2018 to 2019 that show a substantial decrease (from 7 to 7 articles) and an increase from 2019 to 2020 (from 2 to 10); furthermore, there is an increase from 2020 to 2021 (from 10 to 13).

The search outcomes discovered no articles on predictions of Cardiovascular Diseases via DL techniques prior to 2018. That might be clarified by the restricted admittance to open access data sets and the debatably evolving of DL approaches prior 2018. The amount of research papers in this case reduced from 13 articles in the year of 2021 to 4 articles in the year 2022.

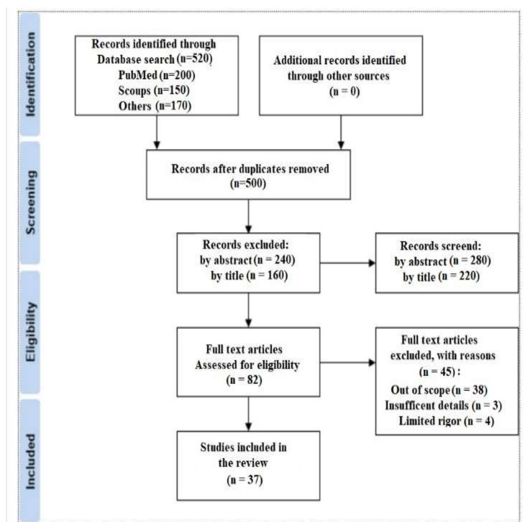

Figure 1: PRISMA Flow Diagram (identification of studies via databases)

Table 2: List of chosen articles for the SLR with their areas of focus

#	Studies	Objective	#	Studies	Objective
1	[1]	Classification of heart sound signals	20	[20]	Heart disease detection
2	[2]	Classification of heart sound signals	21	[21]	Classification of heart sound signals
3	[3]	Classification of heart sound signals	22	[22]	Heart disease detection
4	[4]	Phonocardiogram signals classification	23	[23]	Classification of heart sound signals
5	[5]	Diagnosis of heart sound	24	[24]	Diagnosis of heart sound
6	[6]	Automatic heart sound classification	25	[25]	Classification of heart sound signals
7	[7]	Diagnosis of heart sound	26	[26]	Heart disease detection

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

8	[8]	Improving cardiac auscultation accuracy	27	[27]	Heart sound analysis
9	[9]	Classification of heart sound signals	28	[28]	Heart disease detection
10	[10]	Diagnosis of heart sound	29	[29]	Cardiovascular disease classification
11	[11]	Classification of heart sound signals	30	[30]	Heart disease detection
12	[12]	Heart disease detection	31	[31]	Diagnosis of heart sound
13	[13]	Classification of heart sound signals	32	[32]	Heart disease detection
14	[14]	Diagnosis of heart sound	33	[33]	Heart disease detection
15	[15]	Classification of heart sound signals	34	[34]	Diagnosis of heart sound
16	[16]	Multi-class heart sounds	35	[35]	Cardiovascular disease prediction
17	[17]	Recognition of Heart Failure	36	[36]	Classification of heart sound signals
18	[18]	Diagnosis of heart sound	37	[37]	Heart disease detection
19	[19]	Heart disease detection			

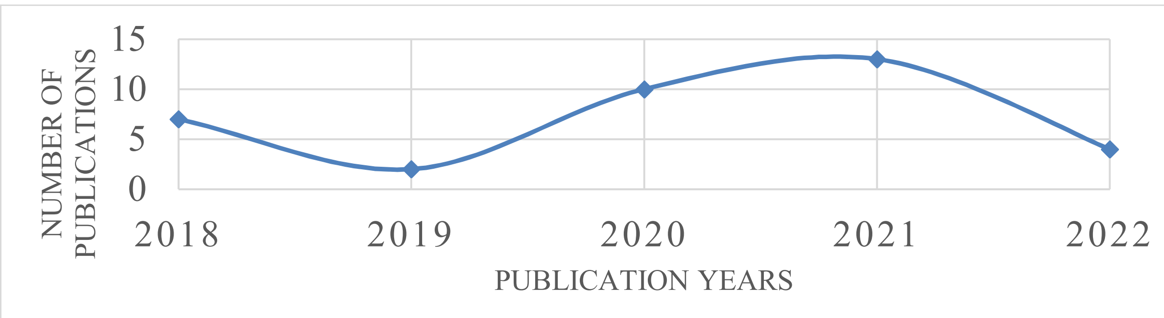

Figure 2: Number of Publications on CVD using DL and ML Techniques between 2018 and 2022

3.1 Algorithms

The SLR considered all stated procedures used in preceding studies. Table 3 and Figure 3 display a set of procedures used in the 37 articles and the amount of articles that embraced these techniques. The SLR showed that the top commonly employed ML procedure was Support Vector Machine (n=11, 28%), followed by K-Nearest Neighbors (KNN) (n=5, 12%). On the other hand, CNN showed most frequently used deep learning methods (n=16, 40%), followed by Recurrent Neural Network (n=6, 15%), ANN (n=2, 5%). 7% of the studies engaged multiple algorithms. One study engaged (CNN and RNN) supervised-learning DL and ML procedures and assessed their performance [28]. 97% of the articles used supervised-learning algorithms of classification to handle the attributes. Figure 3, shows that the methods SVM, ANN, KNN, CNN, RNN reserved their reputation. CNN has got more courtesy in this area than the others. From the years 2018 to 2022, at least one CNN article was published. Another frequently used method is Support Vector Machine (SVM).

Even though RNN have proved great analytical influence, they have documented few apps in CVDs field. The years (2018, 2021, and 2022) documented an increase in the use of assembling and boosting methods from (one to three) papers and no papers in 2019. Furthermore, while no paper used the reinforcement-learning nor semi-supervised approaches, simply one paper employed unsupervised-learning procedure (K-Means) for CVD prediction, though the procedure achieved fine. RF, DT, LR and Boosting Ensemble have recorded zero number of applications during the period 2018-2022.

Table 3: No. of DL and ML methods used in previous works

Algorithms used	No. of Research Papers
ANN	2
CNN	16
KNN	5
SVM	11
RNN	6

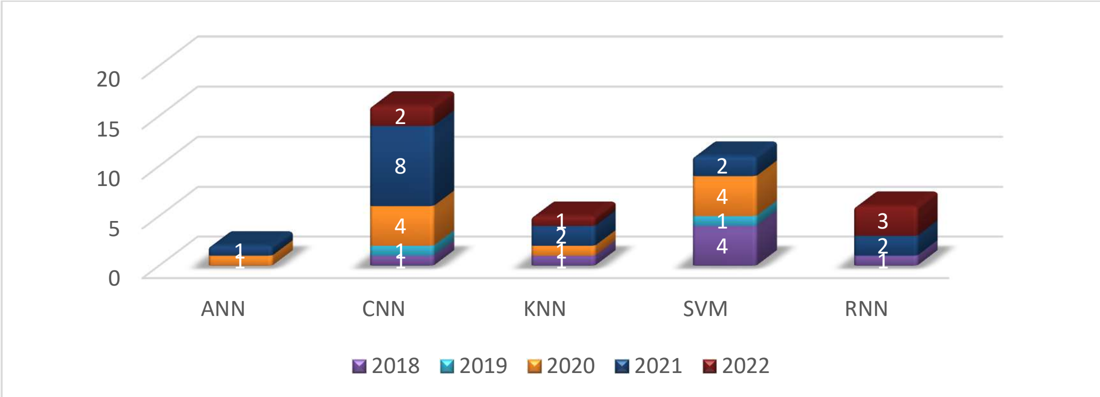

Figure 3: Shows the Yearly usage Frequencies of Popular DL and ML Methods for CVD Prediction.

3.2 Algorithm Performance Metrics

One of the most crucial processes in the prediction process of ML is model assessment. Performance of an algorithm can be measured by means of a selection of metrics. Throughout the training process, measurements are frequently done using hidden examples. The algorithm metrics measurement performance applied in the included studies were "Precision, Specificity, Accuracy, Recall/Sensitivity, and F1-Score". The SLR discovered that the top used metrics assessment in prediction of CVD are "Accuracy, followed by Recall/Sensitivity, F1-Score, Precision, and Specificity". Table 4 shows that all 37 papers stated the accuracy rate used in these models for the prediction of CVD. Figure 4 outlines the percentages of usage of the metrics in the examined articles. This study concentrated on analyzing algorithms performance of the used in the aforementioned studies. But, because it is not suitable to compare the efficiency of two procedures or systems directly if they were assessed on dissimilar datasets [9], the evaluation of the best-performing procedures is according to the same datasets that were used. Illustration on the top method used for assessment (i.e. Accuracy), the top performance algorithms were identified according to the mean values of the accuracy rate of the algorithms gotten from the 37 papers. For example study scope and health data vary greatly among CVD prediction studies, a comparison can only be made after a consistent benchmark on the dataset and scope are established. Therefore, only studies that implemented multiple machine learning methods were carefully chosen for the comparison of the same data and CVD prediction. The authors determined the algorithm with top performance for the same situation by relating the accuracy mean scores of the algorithms that employed the same datasets. Table 5 lists the algorithms used on the different datasets and the calculated mean scores with regard to Recall/sensitivity, specificity, precision, Accuracy, and F1-score. The current study draws on the calculated mean scores of the 37 articles to rank the most performing methods.

Thus, the higher the accuracy of an algorithm, the greater is the chance of creating perfect predictions. From Figure 5, PhysioNet/CinC Challenge Dataset is the most common one used and ANN model scored (97.89%) which is the highest accuracy rate prediction, after that CNN (97.35%), SVM (94.03%), and KNN (85.70%). For the PhysioNet/CinC Challenge 2016 Dataset, the RNN (94.82%), followed by CNN (94.82%) and SVM (88.60) models attained the top analytical accuracy percentage, as can be seen in Figure 6. For the GitHub PCG Dataset, the SVM scored (97.00%) and CNN (85.89%) models achieved the top prediction accuracy as can be seen in Figure 7. For the Pascal Challenge Dataset the top performance achieved by KNN (99.00%) and CNN (96.25%) as shown in Figure 8. As for the Private Dataset KNN (97.00%), CNN (96.30%), SVM (90.18%), RNN (90.11%) and ANN(89.40%) as in Figure 9.

Table 4. Algorithm performance metrics applied in the reviewed studies

Metric	Number of studies
Accuracy	37
F1-Score	9
Precision	9
Recall/Sensitivity	12
Specificity	11

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4
© 2023 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

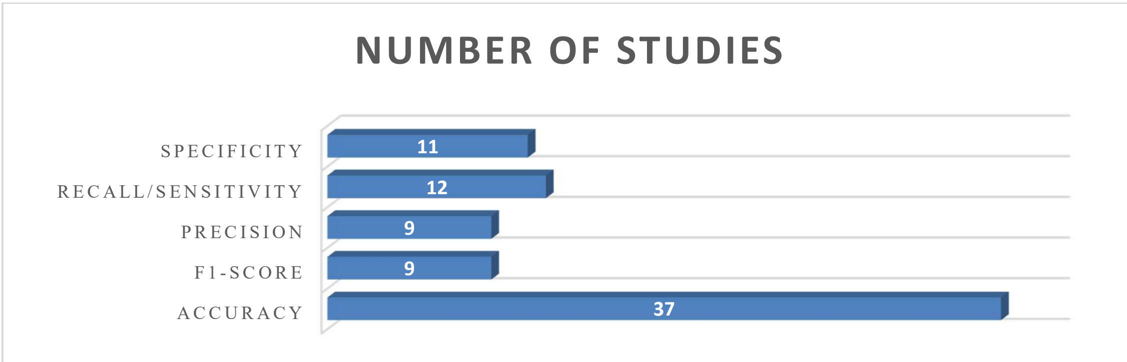

Figure 4: Number of Algorithm Performance Metrics Applied in Reviewed Studies.

Table 5: Algorithm performance based on evaluation metrics and datasets

Dataset	Sample	Algorithms	ACC	F1-Score	Precision	Recall/Sensitivity	Specificity
AptaCDSS-E	66	ANN	94.00	-	-	-	-
		ANN	79.00	-	-	-	-
GitHub PCG	-	CNN	86.57	84.68	96.32	86.52	-
	-	SVM	97.00	-	-	-	-
	-	CNN	85.00	-	-	-	-
Pascal Challenge	176	CNN	96.25	-	-	-	-
		KNN	99.20	-	-	98.57	94.00
		KNN	98.80	-	-	99.99	97.56
PhysioNet/CinC Challenge	200	ANN	97.89	-	-	97.73	98.05
		CNN	98.00	-	-	-	-
		KNN	85.70	85.47	85.30	-	-
		SVM	99.13	-	-	99.00	99.00
		CNN	99.01	-	-	-	-
		SVM	95.40	-	-	-	-
		CNN	95.05	85.59	85.80	-	-
		SVM	93.60	-	-	-	-
		SVM	88.00	-	-	88.00	87.00
PhysioNet/CinC Challenge 2016	3240	CNN	95.00	-	-	-	-
		RNN	96.00	96.00	96.10	-	-
		SVM	90.40	90.20	90.20	90.60	85.50
		CNN	97.89	-	-	-	-
		SVM	86.80	-	-	87.00	72.00
		CNN	96.40	-	-	78.10	87.30
		RNN	96.40	96.09	96.09	-	-
		CNN	90.00	-	-	90.30	90.58

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

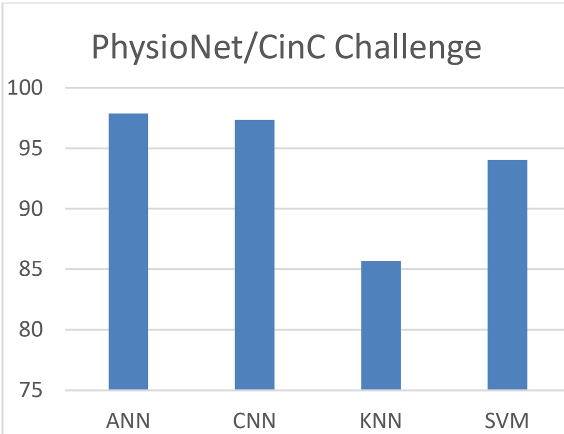

Figure 5: Best Performing Algorithm - PhysioNet/CinC Challenge Dataset for CVD Detection.

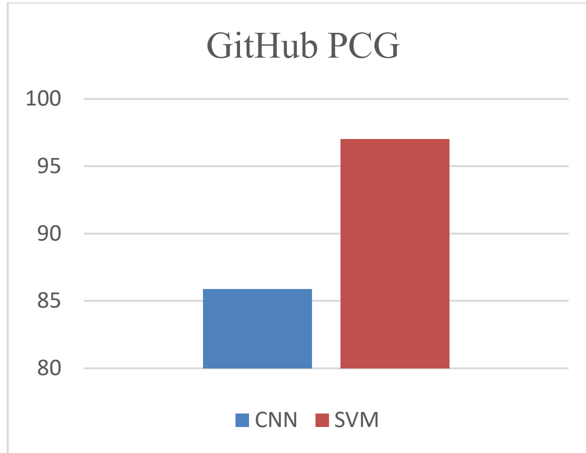

Figure 7: Best Performing Algorithm - GitHub PCG Dataset for CVD Detection.

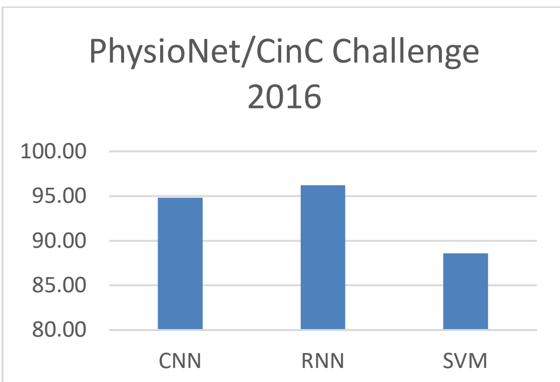

Figure 6: Best Performing Algorithm - PhysioNet/CinC Challenge 2016 Dataset for CVD Detection.

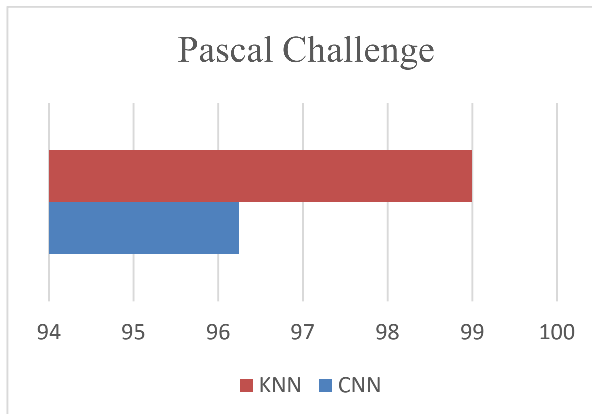

Figure 8: Best Performing Algorithm - Pascal Challenge Dataset for CVD Detection.

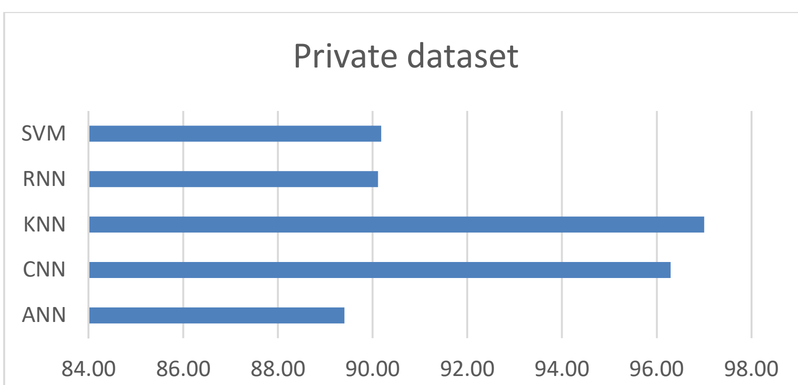

Figure 9: Best Performing Algorithm – Private Dataset for CVD Detection

Table 6: Performance metrics for unpublished primary datasets

Dataset	Size	Algorithms	ACC	F1-Score	Precision	Recall/Sensitivity	Specificity
The interspeech ComParE Heart Sound Shenzhen (HSS)	-	SVM	77.13	-		-	-
The PASCAL Heart Sound Challenge (HSC)	-	CNN	96.00	90.00	90.10	81.00	96.00
University of California at Irvine	-	SVM	85.00	-		-	-
Private datasets	-	SVM	97.53	-		97.00	94.94
	-	ANN	97.20	-		-	-
	-	CNN	93.89	-		-	-
	-	KNN	97.00	-		-	-
	-	RNN	90.11	-		-	-
	-	SVM	95.00	-		-	-
	160	ANN	81.60	-		-	-
	-	CNN	98.70	-		-	-
	-	SVM	78.00	-		-	-
PCG recordings	1000	CNN	98.20	-		-	-
iStethoscope	1636	RNN	94.00	94.00	94.10	-	-
Medical Information Mart for Intensive Care	-	RNN	85.90	85.57	85.20	-	-
Michigan Heart Sounds and Murmur Database (UMHS)	-	CNN	85.00	-		-	-

3.3 Datasets and Data Source

In DL and ML learning, datasets must be trained and validated to test models for accurate predictions of CVD. In the articles reviewed in this study, two key data sets were detected: (1) open-access data sets and (2) un-published key data sets (as in Figure 11). The open access data sets engaged in those articles comprise the AptaCDSS-E, GitHub PCG, Pascal Challenge, PhysioNet/CinC Challenge, PhysioNet/CinC Challenge 2016. The open-access group comprises 25 studies and 5 databases (Table V), but some datasets were used in more than one study like the PhysioNet/CinC Challenge 2016 CVD data set collected from the Kaggle depository. The dataset contained 3240 patients with CVD. According to the homepage of the dataset, most researchers utilize the sound of CVD for the detection of whether a patient has CVD or not. Out of the 37 papers used in this study, 16 papers applied the data set to study the CVD prediction. The second top prevalent data set employed in previous CVD prediction studies was Pascal Challenge CVD dataset. The dataset contains sound files arranged into two groups: Normal, or abnormal. It categorizes patients into Healthy or not healthy. The data set was engaged in three articles and is accessible from the Kaggle Depository.

3.4 Software/Tools used for the CVD Prediction

Deep and machine learning-based approaches are commonly employed for the prediction of CVD. Numerous tools and programming methods were employed for the development of the methods for the CVD predictions. The reviewed studies applied different software for the analyses. These software/tools were classified into programming and data mining software (as in Figure 10). The programming language for deep and machine learning data analysis reported by the articles comprises Python programming platforms like R programming environment, Jupyter Notebook. WEKA, MATLAB, and Minitab are the commonly used data-mining app for CVD predictions in these studies. Of the 37 papers reviewed, 81% (30 papers) used programming languages and 19% employed data-mining tools. An evaluation and comparison of the various models presented the top accuracy of CVD prediction with the programming languages. Numerous models, such as CNN, SVM, and KNN have been tested on the PhysioNet/CinC Challenge 2016 CVD dataset, for example, using python and WEKA. The results showed that RNN, CNN, SVM achieved accuracy rates of (96.20%), (94.82%) and (88.60%) with Python. The assessment outcomes display that the recall, specificity, accuracy values of the various algorithms improved with Python.

Table 7: Software used for dl and ml data analyses

Study	Software/Tool	Algorithms	Accuracy
[1]	Python	CNN	95.00
[2]	R	KNN	97.00
[3]	Python	CNN	98.00
[4]	Python	CNN	97.89
[5]	Python	CNN	98.20
[6]	Python	CNN	85.00
[7]	MATLAB	SVM	90.40
[8]	Python	CNN	96.00
[9]	Python	KNN	99.20
[10]	Python	CNN	86.57
[11]	Python	SVM	86.80
[12]	R	RNN	94.00
[13]	R	SVM	77.13
[14]	Python	CNN	93.89
[15]	MATLAB	ANN	94.00
[16]	Python	CNN	96.25
[17]	Python	RNN	85.90
[18]	MATLAB	RNN	90.11
[19]	Python	KNN	88.70
[20]	Python	SVM	85.00
[21]	R	SVM	97.00
[22]	Python	SVM	99.13
[23]	WEKA	SVM	97.53
[24]	R	SVM	88.00
[25]	Python	CNN	96.40
[26]	Python	RNN	96.40
[27]	R	CNN	99.01
[28]	Python	CNN	98.70
[29]	Python	ANN	97.89
[30]	WEKA	SVM	95.00
[31]	Python	CNN	95.05
[32]	Python	KNN	98.80
[33]	WEKA	SVM	95.40
[34]	Python	CNN	90.00
[35]	MATLAB	KNN	97.00
[36]	Python	SVM	93.60
[37]	R	CNN	85.00

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

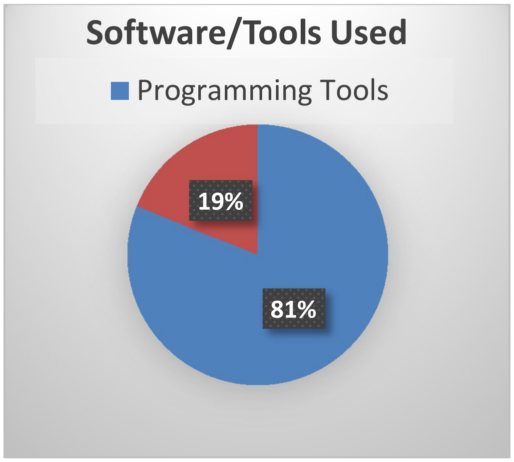

Figure 10: Distribution Analysis of Tools Used.

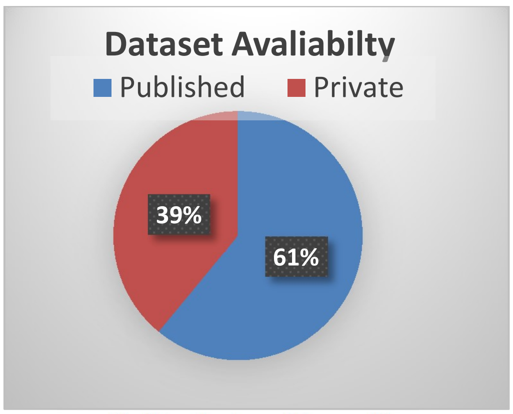

Figure 11: Distribution of Dataset Types used.

4. DISCUSSION

The current study has reviewed 37 articles that use DL and ML methods for the prediction of CVD systematically. By following the questions of the review, it pursued to recognize the DL and ML methods employed for the predictions of CVD, the dataset used, techniques of evaluation used, the software used and top-performing methods for the investigation. The consequence of this analysis high-lights the up-to-date DL and ML methods used in the CVDs predicting gaps, future studies and their performances must be looked out. The outcomes display that preceding articles have concentrated on supervised ML detection based methods for observed methods. The review specified that Convolutional Neural Network, RNN, SVMs, and KNN were the top used methods in predictions of CVD articles, after that Artificial Neural Network. There are inadequate apps of sophisticated and innovative ensemble methods like XGBoost in spite of attaining attractiveness as an ensemble algorithm that has practically confirmed to be an exceedingly efficient method by achieving the top outcomes in many ML oppositions [17]. For the prediction of CVD using DL and ML, no study applied, Adaboost, XGBoost, Bagging, Boosted Decision Tree, have got the least care

This SLR also concentrated on examining the performing of top DL and ML methods for prediction CVD according to the stated assessment methods. The SLR recognized five measurement metrics for performance of model assessment. Those metrics are: "F1-Score, Accuracy, Specificity, Recall/Sensitivity and Precision". This amount is considered sufficient for understanding and likening the study outcomes. As well as other metrics like AUC in upcoming articles is vital. AUC is considered an improved measure of classifiers than the accuracy because it's un-biased environment on the testing data. In regards to the top-performing methods, the outcomes exhibited that hyper methods perform better than a standalone method for prediction of CVD in regards to comparison of accuracy. This outcome opposes the results of [31] since the authors measured the performance of the methods by the AUC metric evaluation. Overall, the analysis prove that the DL and ML Accuracy methods are generally among (0.8 to 0.9+) in the prediction of CVD. This specifies that the ability of the prediction of DL and ML methods are promising in CVD, mainly with KNN, RNN, CNN, and SVM methods. Nonetheless, there may be operational obstructions to similar Accuracy of clinician level. As, there are inadequate circumstances for training and testing of the model. Consequently, additional articles likening DL and ML algorithms and human knowledge are compulsory. Furthermore, the ideal cutoff for accuracy still vague in the surveyed articles. As an example, an AUC score of 0.96 or greater is suggested, but this is not clear with Accuracy.

SLR likewise specified that out of the 37 articles, merely five engaged key medical datasets. It is recommended that data like this be employed in forthcoming articles to uphold the obligation to predict realistic CVD in local settings. This permits us to compare the consequences and sighted the real disadvantages or advantages of the suggested methods. The article likewise proposes that the publication of key clinical data sets and articles will certainly impact forthcoming improvements. The study did not find typical strategies for data splitting. Most articles engaged a cross validation approach and a (60:40, 70:30 or 80:20) dividing approach for the validation and training data sets. Moreover, since the size of most samples of datasets was pretty small, the combined consequences could be unfair. This SLR displays that top articles engaged data-mining tools fewer than programming language, containing Python, R, Google Collab, Jupyter Notebook, etc. Overall, the performance prediction in regards to the accuracy percentage of the methods (i.e., KNN, ANN, and SVM) achieved with the data-mining tools enhanced with Python and WEKA on the identical data set. Nonetheless, the run-time of a known method is likewise critical since if such a system is to be engaged in serious care divisions, a quick judgment needs to be completed.

4.1 Gaps and Future Research Directions

The novel study signifies the first SLR of DL and ML in CVDs predictions. Assuming that disease prediction can aid pull attention to unnecessary involvements, it is vital to identify the up-to-date models of prediction, their performance of prediction, the environment of data sets, and the analysis of technologies. This review is important because it offers an opportunity to enhance these methods. According to the outcomes of the current study, forthcoming investigators must study these gaps:

- A lot more studies employing deep learning methods: similar to Xcpetion, Res-Net, Inception, VGG, Mobile-Net networks are expected.
- Inadequate articles concentrated on Reinforcing Learning and clustering methods.
- Further articles engaging ensemble methods, for instance the Support Vector Machine and Ensemble of Logistic Regression (ELR) are recommended for enhanced prediction.
- Approximately 50% of the involved articles were done in China or the USA. Studies from Asia, Americas (outside the USA) and Africa were inadequate. This may be partially because of the restricted availability of standard structured health data. More articles from the perspective of developing countries are mandatory.
- A main dependence on small sized sample-datasets in the involved articles. Since this may influence the enactment of DL and ML methods, articles with greater data sample sizes are mandatory.
- Involved articles hardly measured performance prediction in regards to AUC, which is recommended to be the top accuracy metric of measurement for classifiers. Forthcoming articles may emphasize on taking AUC as a metric measure of performance.

5. CONCLUSION

While predictions of CVD using DL and ML apps are being commonly investigated, numerous concerns continue to be un-addressed. This research utilized the SLR method to examine up-to-date DL and ML methods used for CVD predictions, assessment methods used and top-performing methods, the data set used, and tools used for the examination. This study showed that a selection of methods can be smeared for predictions of CVD. On the other hand, all methods are a member of a one class which is supervised-learning classification techniques; most articles use published datasets, while fewer studies use key clinical dataset. CNN, KNN, RNN, SVM, and ANN were found to be the best performing methods for prediction of CVD; and programming analysis of data methods such as Python and R were found to yield greater predictive percentage than data-mining tools like MATLAB and MINTAB.

REFERENCES:

[1] Xiao, B., Xu, Y., Bi, X., Zhang, J., & Ma, X. Heart sounds classification using a novel 1-D convolutional neural network with extremely low parameter consumption. *Neurocomputing*, Vol. *392*, 2020, pp. 153-159.

[2] Son, G. Y., & Kwon, S. Classification of heart sound signals using multiple features. *Applied Sciences*, Vol. *8, No.* 12, 2018, 2344.

[3] Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J., & Fan, H. Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. *Neural Networks*, Vol. *130*, 2020, pp. 22-32.

[4] Dhar, P., Dutta, S., & Mukherjee, V. Cross-wavelet assisted convolution neural network (AlexNet) approach for phonocardiogram signals classification. *Biomedical Signal Processing and Control*, Vol. *63*, 2021, 102142.

[5] Oh, S. L., Jahmunah, V., Ooi, C. P., Tan, R. S., Ciaccio, E. J., Yamakawa, T., ... & Acharya, U. R. Classification of heart sound signals using a novel deep WaveNet model. *Computer Methods and Programs in Biomedicine*, Vol. *196*, 2020, 105604.

[6] Li, T., Yin, Y., Ma, K., Zhang, S., & Liu, M. Lightweight end-to-end neural network model for automatic heart sound classification. *Information*, Vol. *12, No.* 2, 2021, 54.

[7] Li, J., Ke, L., & Du, Q. Classification of heart sounds based on the wavelet fractal and twin support vector machine. *Entropy*, Vol. *21, No.* 5, 2019, 472.

[8] Jeong, Y., Kim, J., Kim, D., Kim, J., & Lee, K. Methods for improving deep learning-based cardiac auscultation accuracy: Data augmentation and data generalization. *Applied Sciences*, Vol. *11, No.* 10, 2021, 4544.

[9] Narváez, P., Gutierrez, S., & Percybrooks, W. S. Automatic segmentation and classification of heart sounds using modified empirical wavelet transform and power features. *Applied Sciences*, Vol. *10, No.* 14, 2020, 4791.

[10] Shuvo, S. B., Ali, S. N., Swapnil, S. I., Al-Rakhami, M. S., & Gumaei, A. CardioXNet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings. *IEEE Access*, Vol. *9*, 2021, pp. 36955-36967.

[11] Li, F., Tang, H., Shang, S., Mathiak, K., & Cong, F. Classification of heart sounds using convolutional neural networks. *Applied Sciences*, Vol. *10, No.* 11, 2020, 3956.

[12] Messner, E., Zöhrer, M., & Pernkopf, F. Heart sound segmentation—an event detection approach using deep recurrent neural networks. *IEEE transactions on biomedical engineering*, Vol. *65, No.* 9, 2018, pp. 1964-1974.

[13] Humayun, A. I., Khan, M., Ghaffarzadegan, S., Feng, Z., & Hasan, T. An ensemble of transfer, semi-supervised and supervised learning methods for pathological heart sound classification. *arXiv preprint arXiv:1806.06506,* 2018.

[14] Kui, H., Pan, J., Zong, R., Yang, H., & Wang, W. Heart sound classification based on log Mel-frequency spectral coefficients features and convolutional neural networks. *Biomedical Signal Processing and Control*, Vol. *69*, 2021, 102893.

[15] Bhavya S., Ranjana P. Impact of Deep Learning Algorithms in Cardiovascular Disease Prediction. National Volatiles & Essential Oils, Vol. 8, No. 4, 2021, pp. 4341-4353.

[16] Banerjee, M., & Majhi, S. (, October). Multi-class heart sounds classification using 2D-convolutional neural network. In *2020 5th International conference on computing, communication and security (ICCCS)*, IEEE, 2020, pp. 1-6).

[17] Li, D., Ma, H., Li, W., Zhao, B., Zhao, J., Liu, Y., & Fu, J. KTI-RNN: Recognition of Heart Failure from Clinical Notes. *Tsinghua Science and Technology*, Vol. *28, No.* 1, 2022, pp. 117-130.

[18] Fu, Z., Hong, S., Zhang, R., & Du, S. Artificial-intelligence-enhanced mobile system for cardiovascular health management. *Sensors*, Vol. *21, No.* 3, 2021, 773.

[19] Rath, A., Mishra, D., Panda, G., & Pal, M. Development and assessment of machine learning based heart disease detection using imbalanced heart sound signals. *Biomedical Signal Processing and Control*, Vol. *76*, 2022, 103730.

[20] Kumar, N. K., Sindhu, G. S., Prashanthi, D. K., & Sulthana, A. S. Analysis and prediction of cardiovascular disease using machine learning classifiers. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, 2020, pp. 15-21.

[21] Son, G. Y., & Kwon, S. Classification of heart sound signals using multiple features. *Applied Sciences*, Vol. *8, No.* 12, 2018, 2344.

[22] Deperlioglu, O., Kose, U., Gupta, D., Khanna, A., & Sangaiah, A. K. Diagnosis of heart diseases by a secure internet of health things system based on autoencoder deep neural networks. *Computer Communications*, Vol. *162*, 2020, pp. 31-50.

[23] Nashif, S., Raihan, M. R., Islam, M. R., & Imam, M. H. Heart disease detection by using machine learning algorithms and a real-time cardiovascular health monitoring system. *World Journal of Engineering and Technology*, Vol. *6, No.* 4, 2018, pp. 854-873.

[24] Kathleen H. Miaoa, Julia H. Miaoa. Coronary Heart Disease Diagnosis using Deep Neural Networks. (IJACSA) International Journal of Advanced Computer Science, Vol. 9, No. 3, 2018.

[25] He, Y., Li, W., Zhang, W., Zhang, S., Pi, X., & Liu, H. Research on segmentation and classification of heart sound signals based on deep learning. *Applied Sciences*, Vol. *11, No.* 2, 2021, 651.

[26] Guo, L., Davenport, S., & Peng, Y. Deep CardioSound: An Ensemble Deep Learning Model for Heart Sound MultiLabelling. *arXiv preprint arXiv:2204.07420,* 2022.

[27] Li, F., Liu, M., Zhao, Y., Kong, L., Dong, L., Liu, X., & Hui, M. Feature extraction and classification of heart sound using 1D convolutional neural networks. *EURASIP Journal on Advances in Signal Processing*, Vol. *2019, No.* 1, 2019, pp. 1-11.

[28] Yang, Y., Guo, X. M., Wang, H., & Zheng, Y. N. Deep Learning-Based Heart Sound Analysis for Left Ventricular Diastolic Dysfunction Diagnosis. *Diagnostics*, Vol. *11, No.* 12, 2021, 2349.

[29] Zeng, W., Yuan, J., Yuan, C., Wang, Q., Liu, F., & Wang, Y. A new approach for the detection of abnormal heart sound signals using TQWT, VMD and neural networks. *Artificial Intelligence Review*, Vol. *54, No.* 3, 2021, pp. 1613-1647.

[30] Panicker, S. Use of Machine Learning Techniques in Healthcare: A Brief Review of Cardiovascular Disease Classification. In *2nd International Conference on Communication & Information Processing (ICCIP)*, 2020.

[31] Al-Issa, Y., & Alqudah, A. M. A lightweight hybrid deep learning system for cardiac valvular disease classification. *Scientific Reports*, Vol. *12, No.* 1, 2022, pp. 1-20.

[32] Khan, M. U., Ali, S. Z. E. Z., Ishtiaq, A., Habib, K., Gul, T., & Samer, A. Classification of Multi-Class Cardiovascular Disorders using Ensemble Classifier and Impulsive Domain Analysis. In *2021 Mohammad Ali Jinnah University International Conference on Computing (MAJICC)*, 2021, pp. 1-8). IEEE.

[33] Aziz, S., Khan, M. U., Alhaisoni, M., Akram, T., & Altaf, M. Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through fusion of temporal and cepstral features. *Sensors*, Vol. *20, No.* 13, 2020, 3790.

[34] Alkhodari, M., & Fraiwan, L. Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings. *Computer Methods and Programs in Biomedicine*, Vol. *200*, 2021, 105940.

[35] Ibrahim, N., Jamal, N., Shaabani, M. N. A. H., & Mahadi, L. F. A Comparative Study of Heart Sound Signal Classification Based on Temporal, Spectral and Geometric Features. In *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, IEEE. 2020, pp. 24-29.

[36] Li, P., Hu, Y., & Liu, Z. P. Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods. *Biomedical Signal Processing and Control*, *66*, 2021,102474.

[37] Tang, H., Dai, Z., Jiang, Y., Li, T., & Liu, C. PCG classification using multi domain features and SVM classifier. *BioMed research international*, *2018*.

[38] El-Habil, B. Y., Abu-Naser, S. S. “Global Climate Prediction Using Deep Learning.” Journal of Theoretical and Applied Information Technology, Vol. 100, No. 24, 2022.

[39] Barhoom, A. M. A., Al-Hiealy, M. R. J., Abu-Naser, S. S. “Deep Learning-Xception Algorithm for Upper Bone Abnormalities Classification.”Journal of Theoretical and Applied Information Technology, Vol. 100, No. 23, 2022.

[40] Barhoom, A. M. A., Al-Hiealy, M. R. J., Abu-Naser, S. S. “Bone Abnormalities Detection and Classification Using Deep Learning-Vgg16 Algorithm.” Journal of Theoretical and Applied Information Technology, Vol. 100, No. 20, 2022, pp. 6173-6184.

[41] Abunasser, B. S., Al-Hiealy, M. R. J., Zaqout I. S. and Abu-Naser, S. S. “Breast Cancer Detection and Classification using Deep Learning Xception Algorithm” International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 13. No. 7, 2022, pp.223-228, http://dx.doi.org/10.14569/IJACSA.2022.0130729

[42] Saleh, A., Sukaik, R., Abu-Naser, S.S. Brain tumor classification using deep learning. Proceedings - 2020 International Conference on Assistive and Rehabilitation Technologies, iCareTech 2020, 2020, pp. 131–136, 9328072