Journal of Computational Science and Intelligent Technologies

Online ISSN: 2582-9041

Volume 1, Issue 1
Volume 1, Issue 2
Volume 1, Issue 3

2020

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 1, pages: 01 –08, March 2020

A Novel Intrusion Detection System in WSN using Hybrid Neuro-Fuzzy Filter with Ant Colony Algorithm

[1]Sarah Salaheldin Lutfi, & [2]Mahmoud Lutfi Ahmed
[1]Aysik Consulting Services, Evans GA, USA.
[2]Georgia Southern University, Statesboro, GA 30458, USA

***Corresponding Author: sarahlutfi7@gmail.com*

Received: *02.01.2020,*
Revised: *05.02.2020,*
Accepted: *15.03.2020,*
Published: *30.03.2020*

DOI:
10.53409/mnaa.jcsit1101

Abstract: With the wide application of wireless sensor networks in military and environmental monitoring, security issues have become increasingly prominent. Data exchanged over wireless sensor networks is vulnerable to malicious attacks due to the lack of physical defense equipment. Therefore, corresponding schemes of intrusion detection are urgently needed to defend against such attacks. A new method of intrusion detection using Hybrid Neuro-Fuzzy Filter with Ant Colony Algorithm (HNF-ACA) is proposed in this study, which has been able to map the network status directly into the sensor monitoring data received by base station, accordingly that base station can sense the abnormal changes in network.The hybridized Sugeno-Mamdani based fuzzy interference system is implemented in both the NF filters to obtain more efficient noise removal system. The Modified Mutation Based Ant Colony Algorithm technique improves the accuracy of determining the membership values of input trust values of each node in fuzzy filters. To end, the proposed method was tested on the WSN simulation and the results showed that the intrusion detection method in this work can effectively recognise whether the abnormal data came from a network attack or just a noise than the existing methods.

Keywords: Hybrid Neuro-Fuzzy Filter, Ant Colony Algorithm, fuzzy filters, wireless sensor networks, security issues.

I. INTRODUCTION

Attacks in Wireless Sensor Networks (WSNs) aim in limiting or even eliminating the ability of the network to perform its expected function. WSNs are networks with limited resources and often deployed in uncontrollable environments that an intruder can easily access. WSN attacks target specific network layer's vulnerabilities but normally affect other layers as well. Local sensor activity at multiple sensor network layers should be monitored and evaluated to detect possible malicious intervention. Intrusion detection is the method of analyzing and checking for signs of potential accidents in a computer device or network and also of unauthorized access. Intrusion detection This normally occurs by gathering data from a number of devices and network sources automatically, again and reviewing the data for safety issues.

According to [1], the annual cost from cyber-crimes to the global economy is estimated to be more than $400 billion in 2014. This statistic stresses the importance of developing proper solutions to ensure the safety of information systems. In order to protect information systems against intruders, different preventive solutions have been deployed by organizations. Intrusion detection system, a type of security management system used to gather and analyze information on a computer network or on a website to help build a different attack safeguard mechanism [2]. Also it is recording massive research works on its system using data mining which have been flooded by many researchers recently as it have the potency of classifying and detecting anomalies within a network.

In many real-time applications, there is a need to secure the communication and perform an effective data transmission. Due to attacks and some hackers, various security issues may occur. A lot of researchers were focussed but since the tools such as anti-virus programs and firewalls are not sufficient to provide constant and reliable sources of information. There is a

need to test the IDS system by upgrading the features of the system and provide integrity.

This work focuses on Hybrid Neuro-Fuzzy Filter with Ant Colony Algorithm (HNF-ACA) helps to predict the malicious node with data authentication. This approach is distinct from the standard algorithms for intrusion detection that treat network traffic condition directly. The innovation is that the network status is mapped to a base station's measurements to reflect changes in network status.

The rest of the paper is structured as follows: In section 2 the related work of IDS in WSN is discussed. The proposed IDS detection mechanism using Hybrid Neuro-Fuzzy Filter with Ant Colony Algorithm (HNF-ACA) is described in section 3. The experimental results and discussion is discussed in section 4. The conclusion and future work is given in section 5.

II. RELATED WORK

In literature, there are a few works that aim to combine between anomaly-based approach and hybrid model to benefit from the advantages of both detection techniques. In [3] proposed a mechanism of Intrusion Detection System (IDS) created in a Cluster-based Wireless Sensor Network (CWSN). The method of feature selection is one of the important factors, which affects the performance of IDS. In the field of wireless sensor networks an IDS architecture inspired by the Human Immune System is proposed[4]. However, this intrusion detection systems are too resource-demanding.

A hybrid clustering method is introduced in [5], namely a density-based fuzzy imperialist competitive clustering algorithm (D-FICCA). However, studying a game-based evolutionary algorithm is considered extremely significant. In [6] proposed a lightweight, energy-efficient system, which makes use of mobile agents to detect intrusions based on the energy consumption of the sensor nodes as a metric. A linear regression model is applied to predict the energy consumption. In [7], a game theoretic method is introduced, namely cooperative Game-based Fuzzy Q-learning (G-FQL). G-FQL adopts a combination of both the game theoretic approach and the fuzzy Q-learning algorithm in WSNs. The energy consumption for intrusion detection and prevention should be focused.

In [8] proposed an appropriate probabilistic model which provides the coverage and connectivity in k-sensing detection of a wireless sensor network. However, when random deployment is required, determining the deployment quality becomes challenging. In [9] proposed a Trust Based Adaptive Acknowledgment (TRAACK) Intrusion-Detection System for WSN based on number of active successful deliveries, and Kalman filter used to predict node trust. Based on hybrid models this work is proposing an efficient hybrid system for sensor network. The goal in this research is to study and implement a new model of intrusion detection that combines the advantages of hybrid model in cluster wireless sensor environment, and surpassing other models proposed in the literature.

III. PROPOSED METHODOLOGY

Initially, the fuzzy based Hybrid Neuro-Fuzzy Filter with Ant Colony Algorithm (HNF-ACA) is considered here to evaluate the detection of malicious nodes presented in the wireless sensor nodes. Generally, the malicious node is stated as the node that does not follow the exact behavior, it means there is any malicious attack such as modifying the message, in some terms the packet may be dropped likewise it goes on. Initially, there is a need for establishing the cluster based routing along with the reputation of a node. Then it is necessary to identify the malicious nodes, these parameters are based on the trust recommendation value. If the process is selected then the routing process is made by maintaining the process of each route. It is necessary to update the neighbor node information to source and destination nodes. Finally, for encryption, the RSA (Rivest-Shamir-Adleman) scheme is used to provide the authentication to all nodes based on public and private keys. At last the HNF-ACA is used for classifying the normal and malicious nodes for efficient data transmission in WSN. The architecture diagram of HNF-ACA based IDS is illustrated in Figure 1.

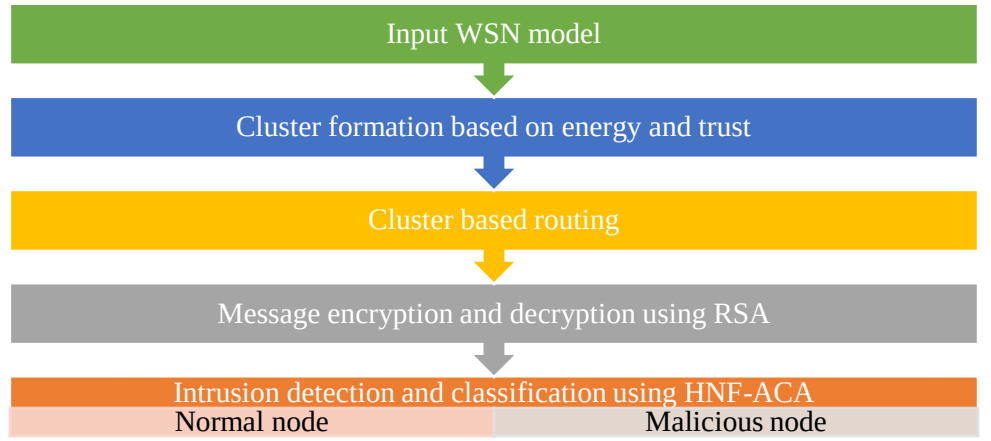

Figure 1. The architecture diagram of HNF-ACA based IDS.

3.1. Cluster Based Routing

The cluster based routing is one of the efficient routing methods, here the nodes are selected for processing as well as sending the data, it is done with the help of high energies. In this work, the Cluster is framed by the nodes that are based on the parameters of residual energy and the trust vector value of reliable nodes. The trust is evaluated based on the successful delivered of packets among source and destination. The routing table of each node consists of watch dog mechanism to monitor the neighbor nodes data and it is necessary to inform the activity. Normally, the routing table is made up of packet id, neighbor coverage range, packet information, cluster member id, link quality, a node to node connectivity etc. The residual energy Res of each node i at time t is calculated as in equation (2), for each and every node to improve the network lifetime [10].

$$Res_i(t) = \frac{[Initial - E_i(t)]}{reg} \qquad (1)$$

where $Initial$ is the initial energy, $E_i(t)$ is the residual energy and reg is the current region of cluster formation. With respect to the high and medium trust vector value the process is determined, the processes of nodes are less energy consumption of the node are considered as a cluster head or cluster members. The trust value $Trust_i$ is calculated as in equation (2),

$$Trust_i = \frac{Deliv_i - packet_i}{recpacket_i} \qquad (2)$$

where, $Deliv_i$ is described as the delivered packets from D_i, $packet_i$ describes the packets sent from D_i to S_i, $recpacket_i$ is describes received packets sent from S_i to D_i, D_i is the destination node and S_i is the source node.

The trust value evaluation is made here to check highest trust value node and find the trusted authenticator. This authenticator indicated as Cluster Head (CH) and other related trust value nodes are selected as cluster member nodes. In case, the selected node is already a member of another cluster region, then a node of high trust value is selected. Based on these CH and members, the cluster region is formed.

Authentication of Message by RSA: The message authentication is processed by RSA algorithm. One of the most popular methods for public key encryption is Ron Rivest, Adi Shamir and Leonard Adleman (RSA) and it is one of the secure publickey encryption methods [12].

3.2. Classification Using HNF-MMACA

To represent the original node representation the authenticate node is used and for misbehavior representation, the malicious node is considered. Based on these two factors the training class is defined. Consider a proposed HNF-MMACA system that has N number of nodes in the input samples. The two classes are considered here in the training example data , i.e., the process has two labels, as well as the proposed system use $K = 2$ like class groups of hidden nodes, and these nodes represents a Gaussian function midpoint with a correlated label. The process is a class grouped; each Gaussian has a different centre but the same label.

Proposed Hybrid Neuro - Fuzzy: Structure of the Hybrid Neuro - Fuzzy, the fuzzy system is classified into Mamdani and Sugeno, which is associated with fuzzy rules. The structures of the two hybrid NF are identical to each other. Each filter is a combination of first order Sugeno and Mamdani fuzzy interference system. Therefore, the general form of the output function is given by equation (3),

$$Y = \sum_{k=1}^{K} \frac{a_k v_k}{\sum_{k=1}^{K} a_k} \; b_k(X) \qquad (3)$$

Where a_k is the degree, the input x matches the rule computed, b_k for the Sugeno model and volume of the output fuzzy set is represented as v_k and the centroid of the output fuzzy set is denoted as N_k. The outcome of both the hybridized NF filter is computed by equation (10), represents the generalized fuzzy interference model which is hybridized the both Sugeno model and Mamdani model. Therefore, the hybridized model is defined as in equation (4)

$$Rule\ k: if\ x_1\ is\ M_1^k;\ x_2\ is\ M_2^k;$$
$$x_3\ is\ M_3^k\ and\ x_4\ is\ M_4^k$$
$$Y\ is\ N_k\left(v_k, b_k(X)\right) \qquad (4)$$

Where the N_k for Mamdani model, or b_k for the Sugeno model. The M_i^k is represented [13] the membership function of the ith input of the kth rule. The essential features are preserved the interpretable capability of the Mamdani model and also enhance the accuracy of the Sugeno model as in equation (5). The hybrid NF outcome fetched into the post processor Y_{post} is truncated to the 8 bit integer values.

The final outcome of the proposed filter Y_{Final} is determined by the average output of the two hybrid NF filters. The function round (), rounds the element x to the nearest integer.

$$Y_{Final} = \text{round}\left(\frac{1}{2} \sum_{\text{post}=1}^{2} Y_{post}\right) \qquad (5)$$

The outcome is fetched into the post processor *post* to produce the final output of the filter. The internal parameters of the hybridized NF filters are optimized using hybrid learning rule to reduce the error. The antecedent and the consequent parameters are optimized using gradient descendent and least mean squares algorithm respectively.

MACA based optimization of Membership values: The membership values are assigned for further optimizing the input to improve the accuracy of the system as it decides the efficiency of IDS [14].

Table 1. Algorithm for optimal membership value Selection using ACA for HNF based IDS

Input: Initial Membership values for nodes, N that represents the number of features (i.e nodes parameters)
Output: Optimized Membership values
Formulate the optimization problem as a graph that is fully connected.
Assign, initial values of membership to the nodes in the network model.
Let number of membership values be *m*;
While best membership value is not optimum do
Assign the parameter Residual energy function is minimum
Spread the membership values on the construction graph.
Update the visited membership values with their error functions.
While optimal membership value is not met do
Initialize ACA
For each ant do
Choose next node by applying state transition rule
Apply step by step pheromone update
End for
Repeat step 10 to 13 until met stopping criterion
Update best solution
Apply offline pheromone update
Check stopping condition not satisfied do
Position each ant in a starting node
For
. Check the Residual energy function of the value
End for
. End while
. Select best membership value in the values m as M_{best};

. Update the global best solution

. End while

. Return $M_{best.}$

Do IDS using HNF

. Calculate the overall feature vector pairs with minimal distance MD

. Calculate the vectors of MD for all classes

. Input can be classified based on the conditions using equation (4)

If the classification is made for all inputs, then the function is terminated,

. Else forward it to the Step 31 process.

. Produce the classification results of normal and malicious node

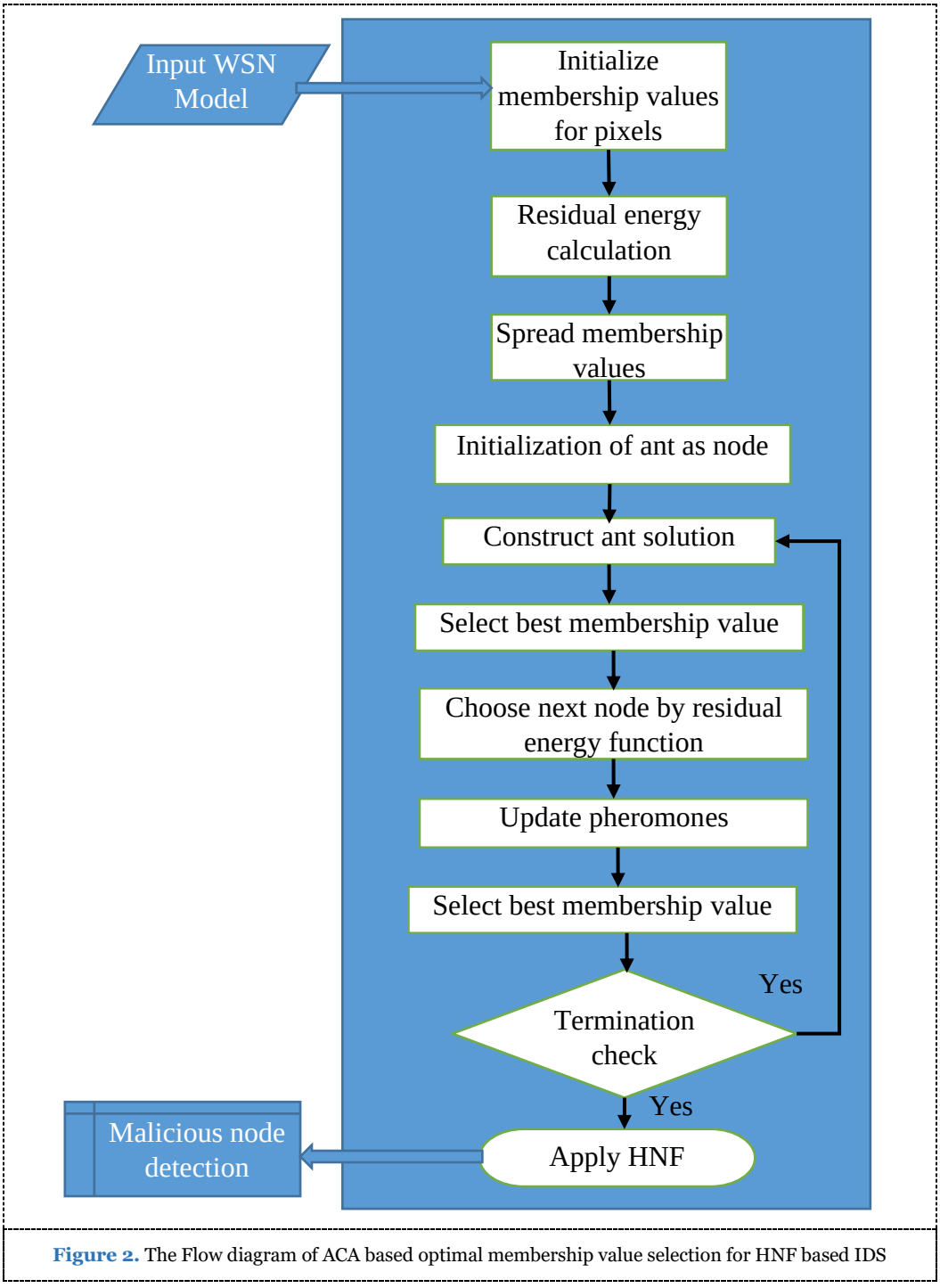

Figure 2. The Flow diagram of ACA based optimal membership value selection for HNF based IDS

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the proposed HNF-ACA performance is carried out with the traditional methods such as D-FICCA [5], G-FQL [7] and TRAACK [9] in presence of malicious node environment in terms of Detection Rate Comparison, communication overhead, End to End delay and network lifetime. The proposed IDS are simulated with Network Simulator tool (NS 2.34). The simulation settings and parameters of the proposed scheme are given in Table 2 [15].

Table 2. Simulation Settings

Number of nodes	101
Area size	1000×1000
Mac	802.11
Radio range	250m
Simulation time	100sec
Packet size	80bytes
Mobility model	Random way point
Protocol	LEACH

4.1. Detetection Rate Comparison

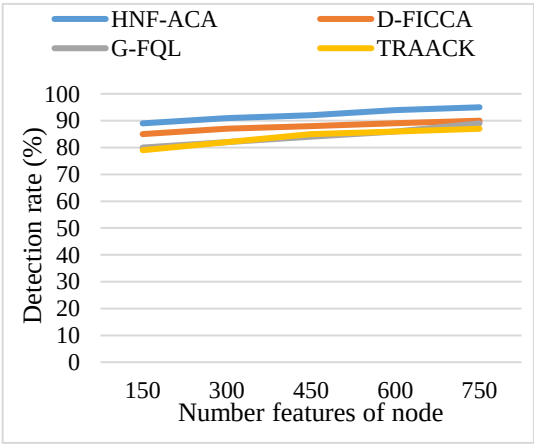

Figure 4. Representation of Detection Rate Comparison

From the figure 4, when increasing the number of features the detection rate is increased twice that of the initial conditions. It is concluded that the detection rate of the proposed HNF-ACA method has improved with rate of 95%. In exiting methods such as D-FICCA, G-FQL and TRAACK attains low rate of 90%, 89% and 87% respectively. The numerical results of detection rate comparison is shown in Table 3.

Table 3. The numerical results of detection rate comparison

No.of features	HNF-ACA	D-FICCA	G-FQL	TRAACK
150	89	85	80	79
300	91	87	82	82
450	92	88	84	85
600	94	89	86	86
750	95	90	89	87

4.2. Communication Overhead Comparison

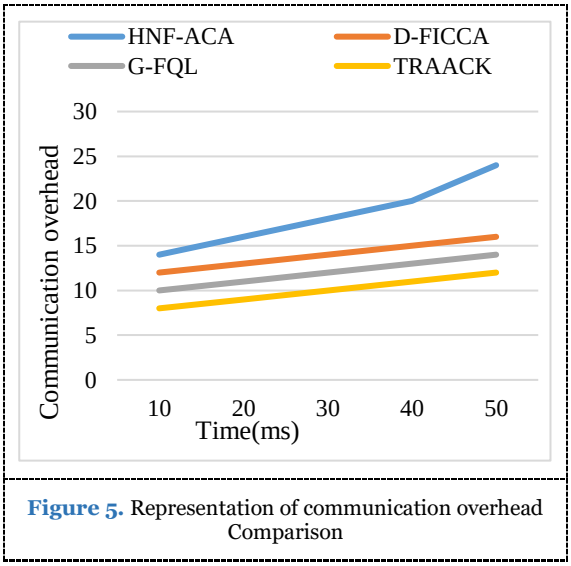

Figure 5. Representation of communication overhead Comparison

Communication overhead is a process of integrating an additional or indirect measure of resource that includes time, memory, bandwidth and a few other resources that are necessary to obtain the objective. From the figure 4, when increasing the time the communication overhead rate is increased. The communication overhead rate of the proposed HNF-ACA method has improved with rate of 24. In exiting methods such as D-FICCA, G-FQL and TRAACK attains low rate of 16, 14 and 12 respectively. The numerical results of communication overhead Comparison is shown in Table 4.

Table 4. The numerical results of communication overhead Comparison

Time (ms)	**HNF-ACA**	**D-FICCA**	**G-FQL**	**TRAACK**
10	14	12	10	8
20	16	13	11	9
30	18	14	12	10
40	20	15	13	11
50	24	16	14	12

4.3. End to end delay Comparison

Figure 6 shows the result of an end to end delay by varying nodes mobility. The obtained delay graph shows that the delay taken by data packets to reach the destination is less for proposed HNF-ACA methodology with 0.2ms when comparing with the traditional D-FICCA, G-FQL and TRAACK attains low rate of 0.24ms, 0.25ms and 0.31ms respectively. The graph proves that the end to end delay is reducing if the mobility of the nodes gets increased.

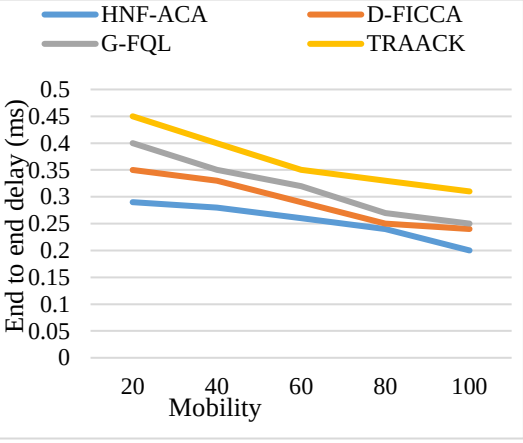

Figure 6. Representation of End to end delay Comparison

Hence, it is summarized that, if nodes get increased then the end to end delay decreased, it is stated as the transmission is effective. The numerical results of End to end delay Comparison is shown in Table 5.

Table 5. The numerical results of End to end delay Comparison

Mobility	HNF-ACA	D-FICCA	G-FQL	TRAACK
20	0.29	0.35	0.4	0.45
40	0.28	0.33	0.35	0.4
60	0.26	0.29	0.32	0.35
80	0.24	0.25	0.27	0.33
100	0.2	0.24	0.25	0.31

4.4. Network Lifetime Comparison

Figure 6 shows the result of Network lifetime by varying nodes time. The obtained delay graph shows that the delay taken by data packets to reach the destination is less for proposed HNF-ACA methodology with 1.21s when comparing with the traditional D-FICCA, G-FQL and TRAACK attains low rate of 1s, 0.9s and 0.8s respectively.

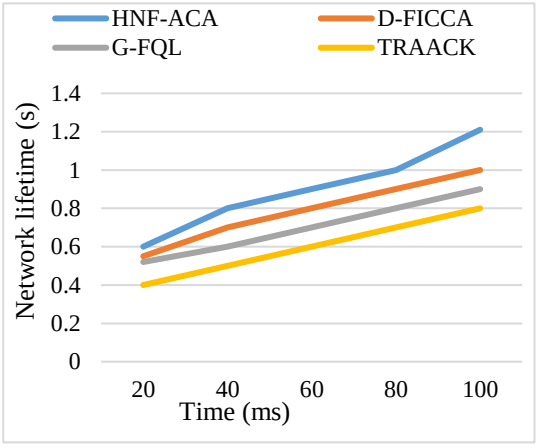

Figure 6. Representation of Network lifetime Comparison

Hence, it is summarized that, the proposed HNF-ACA is highly recommended for IDS in WSN. The numerical results of Network lifetime Comparison is shown in Table 7.

Table 7. The numerical results of Network lifetime Comparison

Time (ms)	HNF-ACA	D-FICCA	G-FQL	TRAACK
20	0.6	0.55	0.52	0.4
40	0.8	0.7	0.6	0.5
60	0.9	0.8	0.7	0.6
80	1	0.9	0.8	0.7
100	1.21	1	0.9	0.8

V. CONCLUSION AND FUTURE WORK

In this work, HNF-ACA based anomaly detection scheme is proposed for large scale sensor networks. The proposed filter hybridizes the Mamdani and Sugeno fuzzy interference model which offers the effective filtering mechanism. Furthermore, the proposed mechanism improves the performance of Sugeno model as well as Mamdani model. The ACA technique greatly improves the membership values to the input. It exploits the stability in their neighborhood information. Initially, the cluster based routing is selected with the trust vector and residual energy of neighbor nodes in the random topology. In this system, the packet transmission from a source node to a destination node is encrypted using the RSA scheme. After that, the identification of malicious nodes is made by using trust recommendation value. Finally, the normal and malicious nodes are classified by using HNF-ACA. An experimental performance result proved that the HNF-ACA provides maximum detection efficiency and high network lifetime, an end to end

delay and overhead is minimized than the existing D-FICCA, G-FQL and TRAACK.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank their present employer for providing support while carrying out this research work.

REFERENCES

[1]. McAfee, Centre for Strategic & International Studies, Estimating the Global Cost of Cybercrime, Technical Report, McAfee, Centre for Strategic & International Studies, 2014.

[2]. Patel A, Qassim Q, Wills C. A survey of intrusion detection and prevention systems. *Inf. Manage. Comput. Secur.* 2010; 18(4): 277-290.

[3]. Wang S S, Yan K Q, Wang S C, Liu C W. An integrated intrusion detection system for cluster-based wireless sensor networks. *Expert Syst. Appl.* 2011; 38(12): 15234-15243.

[4]. Salmon H M, De Farias C M, Loureiro P, Pirmez L, Rossetto S, Rodrigues P H D A, Pirmez R, Delicato F C, da Costa Carmo L F R. 2013. Intrusion detection system for wireless sensor networks using danger theory immune-inspired techniques. *Int. J. Wireless Inf. Networks*. 2013; *20*(1): 39-66.

[5]. Shamshirband S, Amini A, Anuar N B, Kiah M L, Teh Y W, Furnell S. D-FICCA: A density-based fuzzy imperialist competitive clustering algorithm for intrusion detection in wireless sensor networks. *Measurement*. 2014; 55: 212-226.

[6]. Riecker M, Biedermann S, El Bansarkhani R, Hollick M. Lightweight energy consumption-based intrusion detection system for wireless sensor networks. *Int. J. Inf. Secur.* 2015; 14(2): 155-167.

[7]. Shamshirband S, Patel A, Anuar NB, Kiah ML, Abraham A. Cooperative game theoretic approach using fuzzy Q-learning for detecting and preventing intrusions in wireless sensor networks. *Eng. Appl. Artif. Intell.* 2014; 32: 228-241.

[8]. Assad N, Elbhiri B, Faqihi MA, Ouadou M, Aboutajdine D. Efficient deployment quality analysis for intrusion detection in wireless sensor networks. *Wireless Networks*. 2016; 22(3): 991-1006.

[9]. Rajeshkumar G, Valluvan K R. An energy aware trust based intrusion detection system with adaptive acknowledgement for wireless sensor network. *Wireless Pers. Commun.* 2017; 94(4): 1993-2007.

[10]. Khanum S, Usman M, Hussain K, Zafar R, Sher M. Energy-efficient intrusion detection system for wireless sensor network based on MUSK architecture. InHigh Performance Computing and Applications 2010 (pp. 212-217). *Springer*, Berlin, Heidelberg.

[11]. Panda M. Security in wireless sensor networks using cryptographic techniques. *American J. Eng. Res.* 2014; 3(01): 50-56.

[12]. Singh G. A study of encryption algorithms (RSA, DES, 3DES and AES) for information security. *Int. J. Comput. Appl.* 2013; 67: 33-38.

[13]. Kumar PA, Selvakumar S. Detection of distributed denial of service attacks using an ensemble of adaptive and hybrid neuro-fuzzy systems. *Comput. Commun.* 2013; 36(3): 303-19.

[14]. Saidi-Mehrabad M, Dehnavi-Arani S, Evazabadian F, Mahmoodian V. An Ant Colony Algorithm (ACA) for solving the new integrated model of job shop scheduling and conflict-free routing of AGVs. *Comput. Industrial Eng.* 2015; 86: 2-13.

[15]. Nirmaladevi P, Tamilarasi A. Secure Intrusion Detection System for Authentication in Wireless Sensor Networks. Asian J. Res. Social Sci. Humanities. 2016; 6(9): 131-146.

Cite this article as: Sarah S L, Mahmoud L A. Novel Intrusion Detection System in WSN using Hybrid Neuro-Fuzzy Filter with Ant Colony. J. Comput. Sci. Intell. Technol. 2020; 1(1): 1–8. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 1, pages: 09 – 14, March 2020

Classification of Diabetic Retinopathy using Stacked Autoencoder-Based Deep Neural Network

[1]Yasir Eltigani Ali Mustafa, & [2]Bashir Hassan Ismail

[1]*Department of Information Systems, Ahmed Bin Mohamed Military College, Doha, Qatar.*

[2]*Breaking Barriers, United Kingdom.*

[**]***Corresponding Author: yasir@abmmc.edu.qa***

Received: *05.01.2020,*
Revised: *08.02.2020,*
Accepted: *17.03.2020,*
Published: *30.03.2020*

DOI:
10.53409/mnaa.jcsit1102

Abstract: Diagnosis of diabetic retinopathy (DR) via images of colour fundus requires experienced clinicians to determine the presence and importance of a large number of small characteristics. This work proposes and named Adapted Stacked AutoEncoder (ASAE-DNN) a novel deep learning framework for diabetic retinopathy (DR), three hidden layers have been used to extract features and classify them then use a Softmax classification. The models proposed are checked on Messidor's data set, including 800 training images and 150 test images. Exactness, accuracy, time, recall and calculation are assessed for the outcomes of the proposed models. The results of these studies show that the model ASAE-DNN was 97% accurate.

Keywords: Adapted Stacked AutoEncoder, diabetic retinopathy , deep learning and Deep Neural Network.

I. INTRODUCTION

One of the most advanced organs, the human eye is retinal, pupil, iris, lens and optical nerve. Appears as an effective screening method for early detection for eye disease automated retinal image analysis. Retinopathy (DR) and glaucoma, uncontrolled diabetes, can lead to blindness. Diabetes mellitus is elevated blood glucose chronic condition due to either insulin shortage or resistance to insulin [1,2].

The worldwide diabetes prevalence was found at around 425 million in 2017 and is expected to increase to about 630 million in 2045 [3]. In India, about a fifth to a third (57 million) of all people with Diabetes Mellitus (DM) will have retinopathy by 2025. Amongst them, about 5.7 million diabetes sufferers will suffer from extreme retinopathy and will need a laser or surgical operation to maintain vision [4].

Convolutional Neural Networks (CNNs), a deep learning branch, has an outstanding performance of applications such as medical imaging and analysis of images. This work provides the two-stage DR detection method based on the model with retinal images with these motives. At first, the image data of MESSIDOR is used as an input. In the model ASAE-DNN, the DNN-based system is used with DR classification Adapted Stacked Automobile Encoders (ASAE-DNN).

The rest of the paper is written like this. Section 2 deals with the classification of DR related plays. Section 3 describes the methods proposed for defining the DR classification with ASAE-DNN. Section 4 discusses the experimental findings. Section 5 includes the conclusion and prospective work.

II. RELATED WORK

A good computer-aided clinical decision support system was developed to identify retinal images using the neural network and introduced [5]. The literature proposed different methods for the identification of DR. With and without moderate retinopathy, and Bayesian ANN has been educated to distinguish between healthy and diabetic eyes [6]. The proposed methodology to modelling based on *m*-medios [7] was extended and combined with a Gaussian model of a mixture in an ensemble to constructing a hybrid level classification to enhance grade accuracy.

A new automated screening scheme is proposed to support diabetic retinopathy [8] which involves the automatic selection,

screening and classification of colour fundus pictures of diabetic retinopathy that might help diagnose and control diabetic retinopathy.

A novel automated identification with the use of fluid image processing techniques for diabetic retinopathy and maculopatherapy with eye-fundus images [9]. The Machine Learning Bagging Ensemble Classifier (ML-BEC) is designed to classify retinal features for DR disease diagnosis and early detection using computer teaching and ensembling methods [10]. A comprehensive deep learning algorithm based on data was designed and evaluated as a new DR detection diagnostic tool [11].

III. PROPOSED METHODOLOGY

Figure 1 displays the overall design of the Adapted Auto-Encoder stack for Diabetic retinopathy with a Deep neural network. The network uses an image data MESSIDOR as input on the first level and generates a low-level fixed-length function vector from the input. No function engineering is necessary, and the extractor immediately learns in the secret layer, contrary to several traditional DR methods. The ASAE-DNN extracts the data collection with ASAE and classifies the dataset with softmax row. Then, three hidden layers cross the low-level feature vector, and softmax has been used to measure the diabetic retinopathy classification result. During the following supervised training step, the final classification layer (output layer) is applied to practice the final Diabetic Retinopathy predictive models.

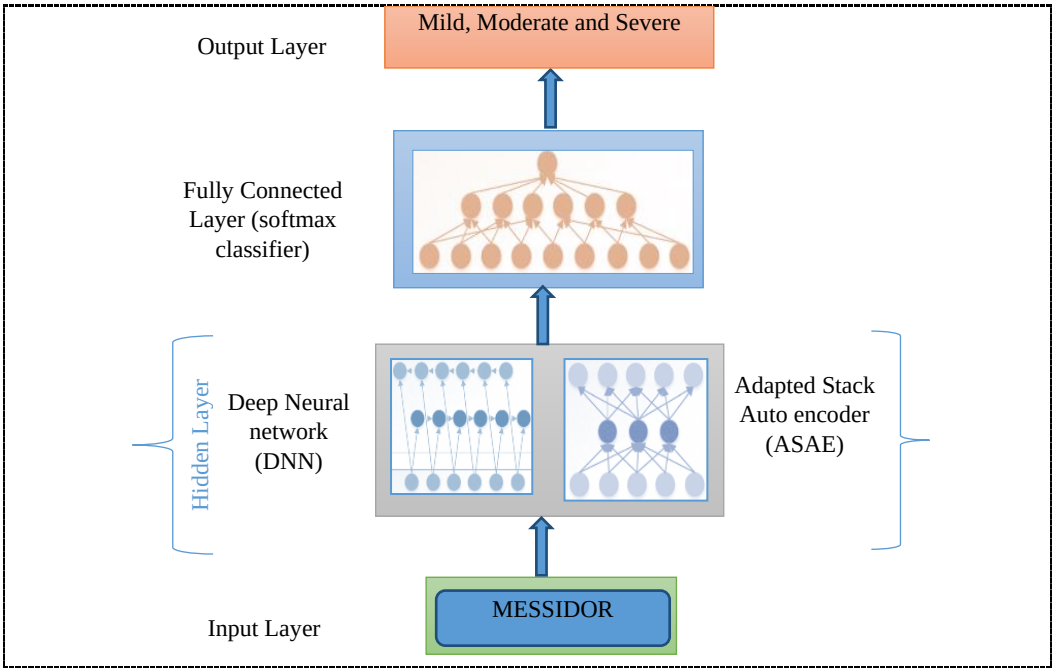

Figure 1. The architecture diagram of classification of DR using ASAE-DNN.

3.1. *Input MESSIDOR dataset*

In this work, to identify DR, a benchmark MESSIDOR dataset was used [12]. This data had approximately 1200 colour fundus images with proper annotation. The images in the dataset were categorized into four categories. Image grading was performed on the existence of microaneurysms, and hemorrhages were allocated to the images. The image without any symptoms indicates healthy retina. The image which has some microaneurysms represents stage 1 (Mild) whereas the image with some microaneurysms as well as hemorrhages denotes stage 2 (Moderate). The images that indicate more microaneurysms, as well as hemorrhages that are placed under stage 3 (Severe).

3.2. *Diabetic Retinopathy classification Using ASAE-DNN model*

Motivated by the fascinating features of deep networks, this work proposes an examination of the entire classification problem using DNN-based structures using adapted stacked autoencoders (ASAE-DNN). The DNN diabetes dataset classifier is built with an adapted autoencoder stacked with three hidden layers of extraction, and the last hidden layer to the classification process is attached to the softmax layer. The output layer gives the

probabilities for a specific record in the diabetic and non-diabetic groups.

Consider an ASAE network with three layers, let $I = 1,2,\dots n$ signifies the raw input of MESSIDOR images also I' is the reconstruction of the image as well as $w^{(l,1)}, w'^{(l,1)}, b^{(l,1)}, b'^{(l,1)}$ that are the weights and bias terms for the lth layer encoder and decoder, individually, $l = 1,2,3$. Primarily, the SAE maps the input image keen on a hidden design hid as signified in equation (1)

$$\sum_{i=1}^{n}\sum_{l=1}^{3} hid_i^{(l)} = \sum_{i=1}^{n}\sum_{l=1}^{3} f\left(I'^{(l)}_i\right) \quad (1)$$

$$f\left(I'^{(l)}_i\right) = s_f\left(w^{(l,1)} I'^{(l)}_i + b^{(l,1)}\right)$$

Then the latent representation $hid(l)$ mapped back into a reconstruction $\mathbb{R}$ from a corrupted version I' as in Eq.(2)

$$\mathbb{R} = g\left(\sum_{i=1}^{n}\sum_{l=1}^{3} hid_i^{(l)}\right)$$

$$g(I^l) = s_g\left(w'^{(l,1)} \textstyle\sum_{i=1}^{n}\sum_{l=1}^{3} hid_i^{(l)} + b'^{(l,1)}\right) \quad (2)$$

where s_f and s_g stand nonlinear activation function. The cost function for SAE well-defined through the reconstruction error $\mathbb{R}(I', \mathbb{R})$ between original input I' and reconstruction $\mathbb{R}$ as in equation (3):

$$J_{ASAE} = \sum_{l=1,2,3} \mathbb{R}\left(I'^{(l)}, g\left(\sum_{i=1}^{n}\sum_{l=1}^{3} f\left(I'^{(l)}_i\right)\right)\right) \quad (3)$$

Where $\mathbb{R}(I', \mathbb{R}) = \|I' - \mathbb{R}\|^2$. Moreover, sparsity constraints term $\mathbb{T}$ is added into hidden units of ASAE to improved learning features, and an additional penalty term is introduced into the objective function in equation (4). Then and there the new cost function of ASAE can be rewritten as in equation (4):

$$J_{MSAE}(\theta) = J_{ASAE} + \mathbb{W} \sum_{j=1\ to\ m} KL(\mathbb{T} \parallel \mathbb{A}_j) \quad (4)$$

The additional penalty terms Kullback–Leibler divergence (KL) in equation (5) is well-defined by the cross-entropy in the middle of $\mathbb{T}$ and $\mathbb{A}_j$.

$$KL(\mathbb{T} \parallel \mathbb{A}_j) = \sum_{j=1}^{m} \mathbb{T} \log \frac{\mathbb{T}}{\mathbb{A}_j} + (1-\mathbb{T}) \log \frac{1-\mathbb{T}}{1-\mathbb{A}_j} \quad (5)$$

$$\mathbb{A}_j = \frac{1}{k}\sum_{i=1}^{k} h_i$$

Here, $\mathbb{T}$ designate the target sparsity level, $\mathbb{A}_j$ denotes the average activation rate for the jth unit, m remains the number of hidden units, $\mathbb{W}$ stands the weight for penalty terms. The sparsity dictions code for each layer represents the input data in the quest for a collection of over-complete base vectors to achieve a better image function. The intrinsic feature of an image is learned by a three-layer scattered coding process, as the sparse representation at pixel level created the feature. Cascading with the softmax classifier is possible to construct a deep network classifier through stacking an autoencoder. Stacked auto encoder may have two or more layers of auto encoders. Let $\{I_1, \dots, I_n\}$ variables be $\{Y_1, \dots, Y_n\}$ input vector given as input to DNN besides the corresponding output class, here $n = 3$. The training method aims to change the DNN parameters to know the input vectors to identify the appropriate output with greater accuracy if the input vectors for the training are assumed. The instructional classification procedure ASAE-DNN is as described in the following:

1. The first autoencoder layer is typically achieved with the original input vector $\{I_1, \dots, I_n\}$ such as the same vector as the destination. This layer tries to recreate the input by removing the $\{\not{f}_{(1,1)}, \dots, \not{f}_{(1,i)}\}$ functions.
2. In addition to yielding the output vector of the first autoencoder layer $\{o_{(1,1)}, \dots, o_{(1,i)}\}$ as an input vector, the second autoencoder layer is equipped by enchanting the output vector as input vector and generates an output vector $\{o_{(2,1)}, \dots, o_{(2,j)}\}$. The second autoencoder layer attempts to recreate the input $o_{(1,i)}; i = 1,2, \dots, C$.
3. The third autoencoder layer is equipped by deciding to take the output vector as the input vector and generates the output vector of the second autoencoder layer $\{c_{(2,1)}, \dots, c_{(2,j)}\}$ as the input vector and generates the output vector $\{c_{(3,1)}, \dots, c_{(3,k)}\}$. The third layer of autoencoders attempts to recreate the o $o_{(2,j)}; j = 1,2, \dots, R$ data.
4. The stacked autoencoder is cascaded via SVM with the softmax classifier layer [13] which will increase the accuracy of DNN classification. This layer is trained through taking output of the third autoencoder layer, $c_{(3,k)}; k = 1,2, \dots, U$ as the input vector and the original class variables $\{Y_{(1)}, \dots, Y_{(N)}\}$ as the query sequence from the training data.
5. Eventually, backpropagation would be used which called fine is tuning, improving the efficiency of the DNN classification. The network is made redundant in some kind of controlled manner with the training data. Once the first hyperplane classification is defined by a back propagation algorithm, the training process is terminated whether it is local or global optimum. The SVM classification hyperplanes are suitable globally by using a comprehensive risk minimisation framework.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed ASAE-DNN is evaluated in this section, and the performance results are compared with existing Neural Network [4], ML-BEC [9] and deep learning

[10] schemes. The performance measurement is done in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (5):

$$Precision = \frac{TP}{FP+TP} \quad (5)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (6):

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

F-measure: this is also referred to as F 1-score, and as in equation (7) is the harmonic mean of precision and recall:

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (7)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (8):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Where true positive (TP) samples are properly classified as no DR, false positive (FP) samples are incorrectly classified as DR, True negative (TN) samples are properly classified as DR, and false negatives (FN) are incorrectly classified as DR.

4.1. Precision Rate comparison

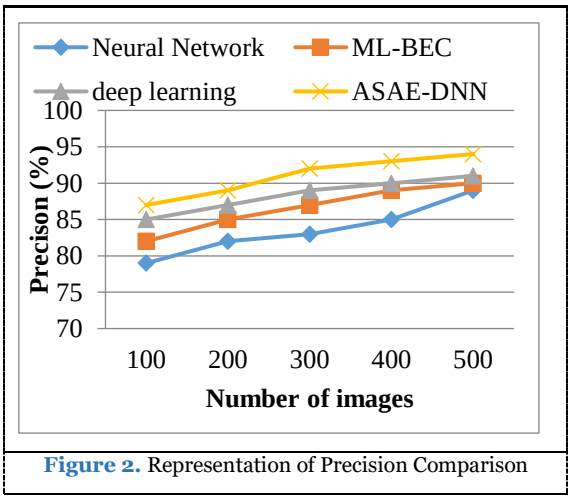

Figure 2. Representation of Precision Comparison

From the above Figure 2, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as Neural Network, ML-BEC and deep learning and ASAE-DNN. When the number of records increases according to the precision value, from this graph, it is learned that the proposed ASAE-DNN offers 94% higher precision than previous methods that yield better results in the classification of CR due to adapted stacked autoencoder.The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

No.of images	Neural Network	ML-BEC	deep learning	ASAE-DNN
100	79	82	85	87
200	82	85	87	89
300	83	87	89	92
400	85	89	90	93
500	89	90	91	94

4.2. Recall comparison

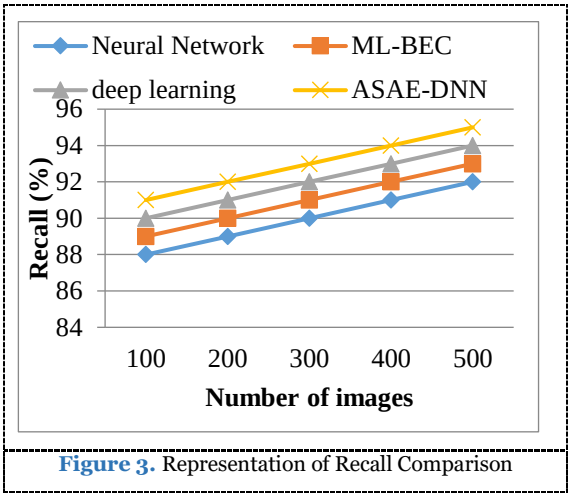

Figure 3. Representation of Recall Comparison

From the above Figure 3, the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as Neural Network, ML-BEC and deep learning. Increasing the number of images often increases the correct value for the recall. Through this graph, it is discovered that the current ASAE-DNN offers recall 95% higher than previous methods. The explanation for this is that the ASAE-DNN extracts the features directly, which will enhance the detection and classification of DR. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

No.of images	Neural Network	ML-BEC	deep learning	ASAE-DNN
100	88	89	90	91
200	89	90	91	92
300	90	91	92	93
400	91	92	93	94
500	92	93	94	95

4.3. *F-measure Rate comparison*

From the above Figure 4, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as Neural Network, ML-BEC and deep learning.

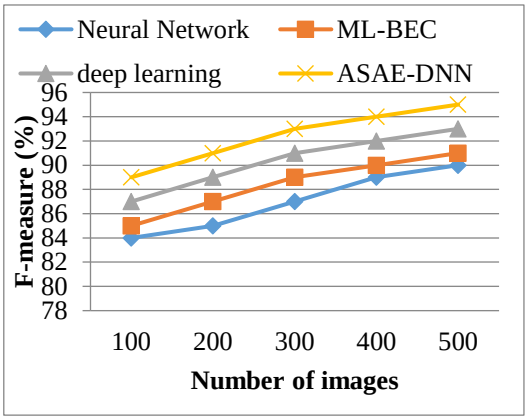

Figure 4. Representation of F-measure Comparison

When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed ASAE-DNN offers 95% higher f-measurement than previous methods. Therefore the proposed ASAE-DNN algorithm is stronger than the current algorithms in terms of better performance of classifying DR. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

No.of images	Neural Network	ML-BEC	deep learning	ASAE-DNN
100	84	85	87	89
200	85	87	89	91
300	87	89	91	93
400	89	90	92	94
500	90	91	93	95

4.4. *Accuracy comparison*

From the above Figure 5, the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as Neural Network, ML-BEC and deep learning and ASAE-DNN. From this graph, it is known that the proposed ASAE-DNN algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better template matching results. This is due to the automatic extraction of the function in the ASAE-DNN algorithm, which increases the DR classification results. The numerical results of Accuracy Comparison is shown in Table 4.

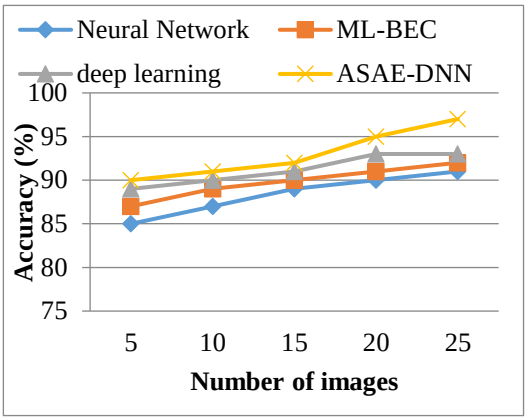

Figure 5. Representation of Accuracy Comparison

Table 4. The numerical results of Accuracy Comparison

No.of images	Neural Network	ML-BEC	deep learning	ASAE-DNN
5	85	87	89	90
10	87	89	90	91
15	89	90	91	92
20	90	91	93	95
25	91	92	93	97

V. CONCLUSION AND FUTURE WORK

This study proposes that ASAE-DNN algorithms use stacked auto-encoders, and that three hidden layers be used for extraction of features accompanied by a softmax classifier classification for DR classification with a high accuracy rate of 97%. It provides a reliable solution for DR detection within a large-scale data set, as well as the results achieved indicate the good efficiency of today's computer-aided model in providing efficient, low-cost and objective DR diagnostics without any need for clinicians to manually examine and grade images. In the future, it may also be important to investigate different types of common patient metadata, such as genetic factors, patient history, duration of diabetes, hemoglobin A1C value, and other clinical data that may influence a patients at risk of retinopathy. Introducing this information to the classification model could give informative correlations to following DR risk factors from outside strictly imaging information, increase thermal diagnostic accuracy.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank their present employer for providing support while carrying out this research work.

REFERENCES

[1]. DeFronzo RA, Ferrannini E, Zimmet P, Alberti KG. International Textbook of Diabetes Mellitus. 4th ed.., Vol. 2. *Oxford UK: Wiley-Blackwell;* 2015.

[2]. International Diabetes Federation. IDF Diabetes Atlas. 8th ed.. International Diabetes Federation; 2017. Available from: http://www.diabetesatlas.org/. [Last accessed on 2019 May 01].

[3]. Balasubramaniyan N, Ganesh Kumar S, Ramesh Babu K, Subitha L. Awareness and practices on eye effects among people with diabetes in rural Tamil Nadu, India. Afr Health Sci 2016;16:210-7.

[4]. Kumar SJ, Madheswaran M. An improved medical decision support system to identify the diabetic retinopathy using fundus images. *Journal of medical systems*. 2012 Dec 1;36(6):3573-81.

[5]. Mookiah MR, Acharya UR, Martis RJ, Chua CK, Lim CM, Ng EY, Laude A. Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach. *Knowledge-based systems*. 2013 Feb 1;39:9-22.

[6]. Somfai GM, Tátrai E, Laurik L, Varga B, Ölvedy V, Jiang H, Wang J, Smiddy WE, Somogyi A, DeBuc DC. Automated classifiers for early detection and diagnosis of retinopathy in diabetic eyes. BMC bioinformatics. 2014 Dec 1;15(1):106.

[7]. Akram MU, Khalid S, Tariq A, Khan SA, Azam F. Detection and classification of retinal lesions for grading of diabetic retinopathy. Computers in biology and medicine. 2014 Feb 1;45:161-71.

[8]. Rahim SS, Jayne C, Palade V, Shuttleworth J. Automatic detection of microaneurysms in colour fundus images for diabetic retinopathy screening. Neural computing and applications. 2016 Jul 1;27(5):1149-64.

[9]. Rahim SS, Palade V, Shuttleworth J, Jayne C. Automatic screening and classification of diabetic retinopathy and maculopathy using fuzzy image processing. Brain informatics. 2016 Dec 1;3(4):249-67.

[10]. Somasundaram SK, Alli P. A machine learning ensemble classifier for early prediction of diabetic retinopathy. Journal of Medical Systems. 2017 Dec 1;41(12):201.

[11]. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. Ophthalmology. 2017 Jul 1;124(7):962-9.

[12]. MESSIDOR dataset: Available at: http://www.adcis.net/en/third-party/messidor/.

[13]. W. Liu, Y. Wen, Yu Z, et al. Large-margin softmax loss for convolutional neural networks, 507–516 (2016)

Cite this article as: Yasir E A M, Bashir H I. New Framework for Anomaly Detection in NSL-KDD Dataset using Hybrid Neuro-Weighted Genetic Algorithm. J. Comput. Sci. Intell. Technol. 2020; 1(1): 09–14. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 1, pages: 15 – 21, March 2020

Classification of Lung Nodules using Improved Residual Convolutional Neural Network

[1]**Afag Salah Eldeen Babiker**

[1]*Department of Computer Science, Faculty of Mathematical Sciences, University of Khartoum, Sudan.*

[**]***Corresponding Author: afagsalah@hotmail.com***

Received: *05.01.2020,*
Revised: *08.02.2020,*
Accepted: *17.03.2020,*
Published: *30.03.2020*

DOI:
10.53409/mnaa.jcsit1103

Abstract: The most common cancer of the lung cannot be ignored and can cause late-health death. Now CT can be used to help clinicians diagnose early-stage lung cancer. In certain cases the diagnosis of lung cancer detection is based on doctors' intuition, which can neglect other patients and cause complications. Deep learning in most other areas of medical diagnosis has proven to be a common and powerful tool. This research is planned for improving the residual evolutionary neural network (IRCNN). These networks apply with some changes to the benign and malignant lung nodule to the CT image classification task. The segmenting of the nodule is performed here by clustering k-means. The LIDC-IDRI database analysed those networks. Experimental findings show that the IRCNN network archived the best performance of lung nodule classification, which findings best among established methods.

Keywords: Deep learning, Improved Residual Convolutional Neural Network,k-means clustering and lung segmentation.

I. INTRODUCTION

Lung cancer is one of the world's most common cancers and causes deaths associated with cancer. It constitutes 13% of all current cancer cases and 19% of worldwide cancer-related deaths. In 2012, it has been estimated that there were 1,8 million new cases of lung cancer [1]. The most common cancer-related and cancer-related mortality in men in India with the largest populations of both male and female Mizoram (age 28.3 and 28.7 per 100 000 population in males and females) [2] is lung cancer that accounts for 6.7 percent of new cases and 9.3 percent of cancer-related deaths in both sexes.

In recent years, conventional AIs have been defeated by neural networks, which are called 'deep learning,' in every crucial task: speaking recognition; images characterisation; and normal, readable sentences producing. In addition to accelerating essential tasks, deep learning enhances device accuracy and CT image recognition and classification efficiency [3]. In this paper, the issue of the classification of benign and malignant is considered and can be solved with k-medium nodule segmentation via the Improved Residual Convolutional Neural Network (IRCNN). The work can be used directly as an input to the complex recovery of data during the extraction and classification of features.

The remaining paper is structured like this. The corresponding works are discussed in section 2. Section 3 outlines the methods suggested for classifying lung nodules. Section 4 addresses the experimental findings obtained. Section 5 contains the conclusion and future work.

II. RELATED WORK

Various approaches for the classification of lung nodules using CT scans have been proposed in the literature. The proposal is made for a new CADe system based on a hierarchical vector (VQ) system [4]. The high-level VQ results in an accurate segmentation of the lungs from the chest volume in comparison with the commonly used simple threshold approach. Based on the Gestalt visual cognition theory, a new lung nodule detection scheme is proposed in [5].

The proposed scheme includes two parts that simulate cognitions of human eyes, for example simple, complete and classified. The nodules were characterized by a gray level co-currency

(GLCM) texture features in the wavelet domain computed by [6] and classified with a radial basis function SVM to classify CT images into two categories: with cancerous lung nodules and without lung nodules.

In [7] classification was focused primarily on the use of supporting vector machines for benign and malignants. The current methods cannot, however, reach great sensitivity and specificities. A method, based on traditional features to compose the contour and texture properties with low frequency curvelet coefficients, is suggested in [8]. Furthermore, the LDA methodology is used for classification labeling. The new intelligent bat algorithm is used to optimize SVM parameters, making it easy and fast. In 9 a new automatic segmentation of the Watershed method lung nodles, multi-size gradient vector flow snakes and the detection method for small lung nodules was introduced using the extracted features and classification. In [10] the effectiveness of DCNNs in the classification of lung nodule malignancy at expert level was assessed.

III. PROPOSED METHODOLOGY

The proposed approach from the Lung Image Database Consortium on the LIDC-IDRI [13] dataset is evaluated in this section. The latter detects lung nodule segmentation from CT images using k-means. For conventional medicine, the complex steps of image extraction function can be minimized by inputting the original image directly into IRCNN. IRCNN based lung nodule classification architecture diagram is illustrated in Figure 1.

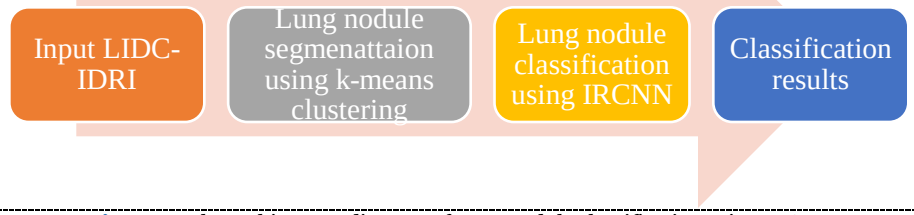

Figure 1. The architecture diagram of Lung nodule classification using IRCNN.

3.1. *Input LIDC-IDRI*

The data set was acquired as a free resource from Lung Image Database Consortium and Image Database Resource Initiative (LIDC / IDRI) that consisted of CT scans of annotated lesions to assess the performance of the proposed nodular-deep system. This database is open to the public, in addition to is used in several studies. An experienced radiologist was asked to mark two types of the lung nodules, such as benign and malignant nodules. Figure 2 represents[10] an example in HRCT scan images of benign and malignant lung nodules collected from the LIDC-IDRI dataset and its boundaries marked by an experienced radiologe. In this study, the 1200 slices were selected from the LIDC-IDRI dataset which contained 2600 equal number of benign and malignant lung nodules.

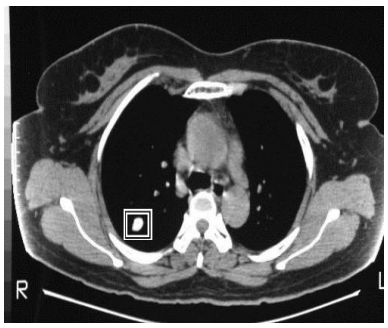

Figure 2. An example of lung nodules of benign.

3.2. *Lung Nodule segmentation Using K-means clustering*

This section used clustering of K-means to detect lung nodule similar to the approach proposed in [11]. Nucleus subspaces are categorized using K-means clustering to use mahalanobis distance metric to cluster the objects into separate clusters. K-means algorithmic rule clustering is projected to find the traditional and anomalous nucleus in the CT picture. In short, the clustering procedure in Kmeans has the following phases:

Step 1: Randomly evaluate the initial centroid point K of the cluster.
Step 2: Find the K-Means clustering objective $C(x)$ to minimize the sum of square distances between all points and the cluster center $C(x) = \sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^j - c_j\right\|^2$, where k denotes the number of clusters, n is the number of CT images, x is the particular image and c is the cluster centroid and $x_i^j - c_j$ is the mahalanobis distance function.
Step 3: Compute the distance of each pixel against the centroid, and group them.
Step 4: Based on each centroid member, measure the New centroid value.
Step 5: Repeat steps 2 and 3 until the current centroid value has not changed from the previous centroid value.

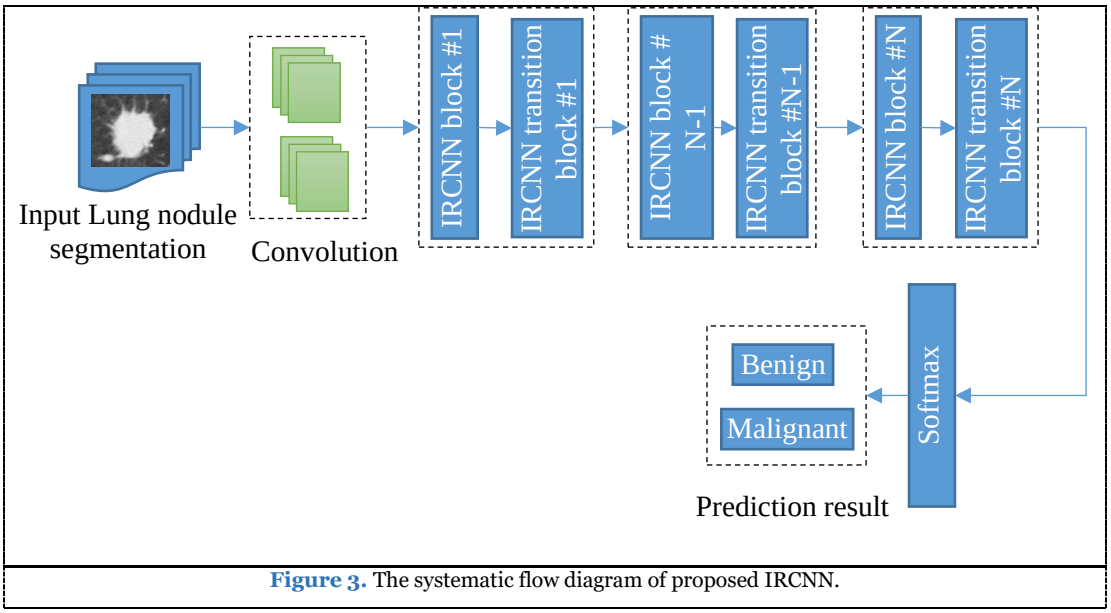

Figure 3. The systematic flow diagram of proposed IRCNN.

3.3. Improved Residual Convolutional Neural Network for lung nodule classification

In this section, the Improved Residual convolution neural network (IRCNN), which utilizes the strength of the Residual convolutional Neural Network (RCNN), the Inception Network, and the Residual Network, is a new Deep learning model. This method increases the precision of Inception-residual network identification with the same number of network parameters. However, this proposed architecture generalizes the Inception Network, the RCNN, and the Residual Network with a significantly improved training precision.

The RCNN block that includes Recurrent Convolution Layers (RCLs), Inception Units, and Residual Units is the most significant part of this proposed architecture. Initially, the inputs are fed into the input layer, at that moment passed through the starting units where RCLs are applied, then finally the outputs of the starting units are added to the RCNN-block inputs. Here, consider the x input sample in the lth layer of the IRCNN-block and a pixel located at (i, j) in an input lung images on the kth feature map (f) in the RCL. Furthermore, assume the output of the network $Y_{ijk}^{l}\,(time)$ is at the time step T. The output can be expressed as in equation (1):

$$Y_{ijk}^{l}\,(T) = \left(w_k^f\right)^{time} * x_l^{f(i,j)}(time) + \left(w_k^{rcl}\right)^{time} * x_l^{rcl(i,j)}(time - 1) + b_k \quad (1)$$

Where $x_l^{f(i,j)}(time)$ and $x_l^{r(i,j)}(time - 1)$ stand the inputs for the standard convolution layers in addition to for the $l\,th$ RCL correspondingly. The w_k^f and w_k^{rcl} values are the weights on behalf of the standard convolutional layer then the RCL of the kth feature map respectively, and b_k is the bias. The outputs o of the starting units for the various size kernels and average pooling layer are determined using equation (2)

$$o = af\left(Y_{ijk}^{l}\,(time)\right) = max(0, Y_{ijk}^{l}\,(time)) \quad (2)$$

Here af is the activation function of standard Rectified Linear Unit (ReLU) of IRCNN. At this time, likewise reconnoitred the performance of this model with the Exponential Linear Unit (ELU) activation function. The outputs o of the inception units for the diverse size kernels at that moment average pooling layer are well-defined as $o_{1\times1}\,(x), o_{3\times3}\,(x)$, and $o_{1\times1}^{p}(x)$ respectively. The

ultimate outputs of IRCNN unit are defined as $\mathbb{Y}(x_l, w_l)$ which can be expressed as

$$\mathbb{Y}(x_l, w_l) = o_{1\times1}(x) \odot o_{3\times3}(x) \odot o^{p}_{1\times1}(x) \quad (3)$$

Where $\odot$ represents the concatenation procedure with reference to the channel or feature map axis of input lung image. The IRCNN-component outputs are then inserted accompanied by the IRCNN-block inputs. The IRCNN-block residual activity can be expressed by equation (4).

$$x_{l+1} = x_l + \mathbb{Y}(x_l, w_l) \quad (4)$$

Where x_{l+1} references to the inputs for the instant subsequent transition block, x_l represents the input image samples of the IRCNN-block, w_l signifies the kernel weights of the l th IRCNN-block, besides $\mathbb{Y}(x_l, w_l)$ characterises the outputs from of l th layer of the IRCNN-unit. Nevertheless the number of feature maps and the elements of the residual unit feature maps are the same as in the IRCNN block. The batch normalization refers to the RCNN-block outputs [12]. The outputs of this IRCNN-block are finally fed to the inputs of the next immediate transition block.

Not the same operations are performed in the transition block, including convolution, pooling, and dropout, depending on where the transition block is located in the network. In the intervening time the size of the input and output features in the IRCNN blocks does not change, it is unbiased a linear projection on the same dimension and the RELU and ELU activation functions add non-linearity. Consequently each convolution layer in the transition block a 0.5 dropout was used here. Eventually, at the end of the architecture, used a Softmax, or normalized exponential function area.

For input lung sample x, weight vector w, and k distinct linear functions, the Softmax operation can be well-defined for the $i\ th$ class as in equation (5):

$$P(o = i|x) = \frac{e^{x^T} w_i}{\sum_{k=1}^{K} e^{x^T} w_k} \quad (5)$$

This proposed IRCNN model was studied and compared across different models by means of a set of experiments on different benchmark datasets. Figure 3 gives a systematic flow diagram for the proposed IRCNN.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed IRCNN is evaluated in this section, and the performance results are compared with existing SVM [7], SVM-LDA [9] and DCNN [10] image compression schemes. The Lung Image Database Consortium image database (LIDC-IDRI) in real time consists of diagnostic and lung cancer screening thoracic computed tomography (CT) scans of annotated marked-up lesions. Seven academic centers and eight medical imaging companies partnered to create this collection of data that contains 1018 cases. That subject contains images from a clinical thoracic CT scan and an accompanying XML file that documents the findings of four experienced thoracic radiologists conducting a two-phase image annotation procedure. The figures given below show that the device proposed has achieved better performance in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (6):

$$Precision = \frac{TP}{FP+TP} \quad (6)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (7):

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

F-measure: this is also referred to as F 1-score, and as in equation (8) is the harmonic mean of precision and recall:

(8):

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (8)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (9):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

Where true positive (TP) samples are properly classified as natural, false positive (FP) samples are incorrectly classified as irregular, True negative (TN) samples are properly classified as irregular, and false negatives (FN) are incorrectly classified as natural.

4.1. Precision Rate comparison

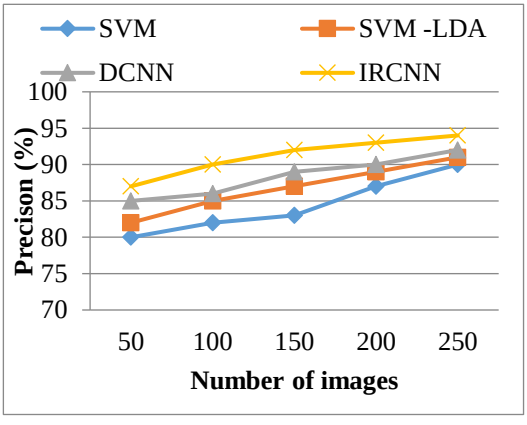

Figure 4. Representation of Precision Comparison

From the above Figure 4, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as SVM, SVM-LDA, DCNN, and IRCNN. When the number of records increases according to the precision value. From this graph, it is learned that the proposed IRCNN offers 94% higher precision than previous methods that yield better results in the classification of lung nodules due to prior segmentation of the lung nodules using k-means technique. The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

No.of images	SVM	SVM - LDA	DCNN	IRCNN
50	80	82	85	87
100	82	85	86	90
150	83	87	89	92
200	87	89	90	93
250	90	91	92	94

4.2. Recall comparison

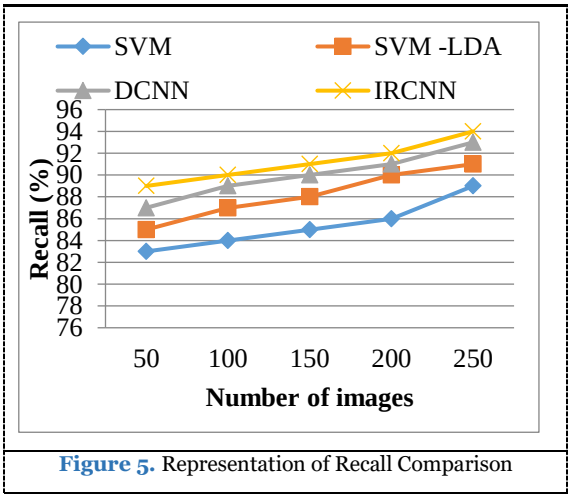

Figure 5. Representation of Recall Comparison

From the above Figure 5 the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as SVM, SVM-LDA, DCNN, and IRCNN. Increasing the number of photographs often increases the correct value for the recall. Through this graph, it is discovered that the current IRCNN offers recall 94% higher than previous methods. The explanation for this is that the IRCNN extracts the features directly which will enhance the lung nodule classification tests. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

No.of images	SVM	SVM - LDA	DCNN	IRCNN
50	83	85	87	89
100	84	87	89	90
150	85	88	90	91
200	86	90	91	92
250	89	91	93	94

4.3. F-measure Rate comparison

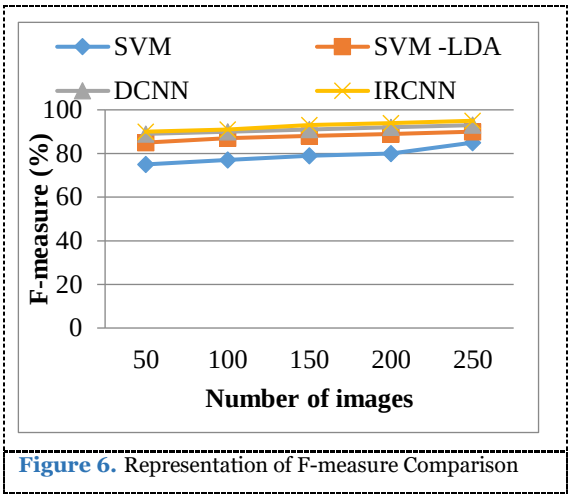

Figure 6. Representation of F-measure Comparison

From the above Figure 6, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as SVM, SVM-LDA, DCNN, and IRCNN. When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed IRCNN offers 95% higher f-measurement than previous methods. Therefore the proposed IRCNN algorithm is stronger than the current algorithms in terms of better performance of classifying lung nodules. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

No.of images	SVM	SVM - LDA	DCNN	IRCNN
50	75	85	89	90
100	77	87	90	91
150	79	88	91	93
200	80	89	92	94
250	85	90	93	95

4.4. *Accuracy comparison*

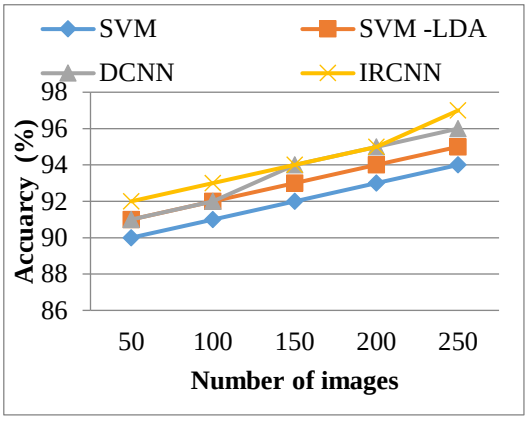

Figure 7. Representation of Accuracy Comparison

From the above Figure 7 the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as SVM, SVM-LDA, DCNN, and IRCNN. From this graph it is known that the proposed IRCNN algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better template matching results. This is due to the automatic extraction of the function in the IRCNN classification algorithm, which increases the classification precision resulting in lung nodules. The numerical results of Accuracy Comparison is shown in Table 4.

Table 4. The numerical results of Accuracy Comparison

No.of images	SVM	SVM -LDA	DCNN	IRCNN
50	90	91	91	92
100	91	92	92	93
150	92	93	94	94
200	93	94	95	95
250	94	95	96	97

V. Conclusion and future work

A classification of the lung nodule is proposed in this work, based on deep learning. This detection scheme can avoid the extraction of candidates and can be less scale based. The k-means algorithm is used for segmentation of the lung nodule, too. Since IRCNN does not recognize anatomical features, several regions of the lung nodule occur in the results of the detection. Experimental results show that most nodules with a high accuarcy rate of 97% can be detected by the designed IRCNN. There are other problems that still need attention with this, such as increasing algorithm sensitivity, reducing the number of false positives, enhancing and optimizing the detection of algorithms of different types of nodules with various sizes and shapes. Based on this analysis, further research is needed to develop current techniques, and new algorithms are needed with machine learning methods to overcome the identified drawbacks.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank their present employer for providing support while carrying out this research work.

REFERENCES

[1]. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. Lyon, France: International Agency for Research on Cancer; 2013. [accessed on January 21, 2014]. GLOBOCAN 2012 v1.0, *Cancer Incidence and Mortality Worldwide*: IARC CancerBase No. 11.

[2]. Indian Council of Medical Research; 2013. [accessed on January 21, 2014]. National Cancer Registry Programme. Three Year Report of Population Based Cancer Registries: 2009-2011.

[3]. Ravì D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, Yang GZ. Deep learning for health informatics. *IEEE journal of biomedical and health informatics*. 2016 Dec 29;21(1):4-21.

[4]. Han H, Li L, Han F, Song B, Moore W, Liang Z. Fast and adaptive detection of pulmonary nodules in thoracic CT images using a hierarchical vector quantization scheme. *IEEE journal of biomedical and health informatics*. 2014 Jun 4;19(2):648-59.

[5]. Qiu S, Wen D, Cui Y, Feng J. Lung nodules detection in CT images using Gestalt-based algorithm. *Chinese Journal of Electronics*. 2016 Jul 1;25(4):711-8.

[6]. Orozco HM, Villegas OO, Sánchez VG, Domínguez HD, Alfaro MD. Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. *Biomedical engineering online*. 2015 Dec 1;14(1):9.

[7]. Dhara AK, Mukhopadhyay S, Dutta A, Garg M, Khandelwal N. A combination of shape and texture features for classification of pulmonary nodules in lung CT images. *Journal of digital imaging*. 2016 Aug 1;29(4):466-75.

[8]. Qiao Z, Kewen X, Panpan W, Wang H. Lung nodule classification using curvelet transform, LDA algorithm and BAT-SVM algorithm. *Pattern Recognition and Image Analysis*. 2017 Oct 1;27(4):855-62.

[9]. Yoshino Y, Miyajima T, Lu H, Tan J, Kim H, Murakami S, Aoki T, Tachibana R, Hirano Y, Kido S. Automatic classification of lung nodules on MDCT images with the temporal subtraction technique. International *Journal of Computer Assisted Radiology and Surgery*. 2017 Oct 1;12(10):1789-98.

[10]. Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks. *International journal of computer assisted radiology and surgery*. 2017 Oct 1;12(10):1799-808.

[11]. Chitade AZ, Katiyar SK. Colour based image segmentation using k-means clustering. *International Journal of Engineering Science and Technology*. 2010;2(10):5319-25.

[12]. Chen H, Zhang Y, Kalra MK, Lin F, Chen Y, Liao P, Zhou J, Wang G. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE transactions on medical imaging*. 2017 Jun 13;36(12):2524-35.

Cite this article as: Afag S E B. Classification of Lung Nodules using Improved Residual Convolutional Neural Network. J. Comput. Sci. Intell. Technol. 2020; 1(1): 15-21. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 1, pages: 22 –28, March 2020

Hybrid Convolutional Neural Network with PSO Based Severe Dengue Prognosis Method in Human Genome Data

[1]**Mohammed Mustafa, [2]Rihab Eltayeb Ahmed, & [3]Sarah Mustafa Eljack**

[1]*Department of Computer Science, University of Tabuk, Tabuk City, Saudi Arabia.*

[2]*Faculty of Computers and information Technology, University of Tabuk, Tabuk City, Saudi Arabia.*

[3]*Majmaah University, Department of Computer Science and Information College of Science Al Zulfi, Saudi Arabia.*

[**]***Corresponding Author: mmustafa@ut.edu.sa***

Received: *05.01.2020,*
Revised: *08.02.2020,*
Accepted: *17.03.2020,*
Published: *30.03.2020*

DOI:
10.53409/mnaa.jcsit1104

Abstract: Dengue is one of the most significant diseases transmitted by arthropods in the world. Dengue phenotypes are focused on documented inaccuracies in the laboratory and clinical studies. In countries with a high incidence of this disease, early diagnosis of dengue is still a concern for public health. Deep learning has been developed as a highly versatile and accurate methodology for classification and regression, which requires small adjustment, interpretable results, and the prediction of risk for complex diseases. This work is motivated by the inclusion of the Particle Swarm Optimization (PSO) algorithm for the fine-tuning of the model's parameters in the convolutional neural network (CNN). The use of this PSO was used to forecast patients with extreme dengue, and to refine the input weight vector and CNN parameters to achieve anticipated precision, and to prevent premature convergence towards local optimum conditions.

Keywords: Deep learning, Particle Swarm Optimization, convolutional neural network and Dengue.

I. INTRODUCTION

Until September this year, dengue has claimed the lives of more than 80 people and affected about 40,000 people across the world. According to statistics under the Ministry of Health's National Vector Borne Disease Control Program (NVBDCP), 83 people have died until 30 September, while 40,868 have been affected. Last year, 325 people were killed, and some 1, 88,401 people were affected by the mosquito-borne tropical disease. Dengue had claimed 35 lives in Kerala and 3,660 had been affected by the State until 30 September 2018. In Maharashtra, 18 people have died and 4,667 have suffered from the disease, according to data [1].

Early forecast factors and algorithms for moderate or severe Dengue prediction were investigated in previous studies. Nevertheless, it is difficult to draw specific conclusions in the case populations (adults vs. infants, hospitalized vs. outpatients) or clinical concepts of "serious" dengue [2] [3]. However, the positive and negative forecast values of the prognostic tool were not reported in previous prognostic studies [4]. In that sense, a practical, forecasting method using CNN-PSO to allow the early prognosis for severe dengue was developed in this current study.

This is the structure of the remaining paper. Section 2 addresses the related plays. The proposed methods of describing the detection of severe dengue using CNN-PSO are defined in Section 3. The experimental findings obtained in Section 4 are discussed. The conclusion and future work are contained in Section 5.

II. RELATED WORK

Specific methods have been suggested in the literature to identify lung nodules using CT dengue. A non-invasive precise diagnostic device was built in [5], which could allow doctors to assess the rates of risk of dengue patients and thus make the right choice. The adaptive neuro-fuzzy inference system (ANFIS) was used to construct the diagnostic model with observations for bioelectronic impedance analysis, with the symptoms or signs provided to dengue patients.

This study [6] aimed to develop a non-invasive diagnostic system to help doctors

classify the risk of dengue-related people using the Levenberg / Marquardt and Scaled Conjugate Gradient algorithms Multilayer Perceptron Neural Network (MPNN) models.

The new Arbovirus Dengue Diagnosis Fusion Architecture was introduced in [7]. The software incorporates characteristics of platelets and CBR to promote the evaluation of the health system. In [8], the early diagnosis of dengue fever was based on a hybrid Artificial Intelligence System, Adaptive Neuro-Fuzzy Inference System (ANFIS). Early dengue disease clinical signs are unspecific and overlap the other infectious diseases.

In [9], attempts have been made to establish an earlier diagnostic model of dengue fever based on PSO-ANN. The PSO methodology used in the proposed model is to optimize ANN process weight and biais parameters. In [10] proposed a new artificial intelligence-based methodology which predicts diagnostics in real time, by using an alternative decision tree approach that promotes highly accurate rules for the generation of false-positive and false-negative diagnosis.

III. PROPOSED METHODOLOGY

The CNN-PSO is theoretical method is used to predict genome data for dengue. The features are entered as CNN-PSO input following pre-processing of the sample dataset and further prediction of severe dengue is done effectively. Figure 1. Illustrates the architectural diagram for the proposed CNN-PSO dengue forecast.

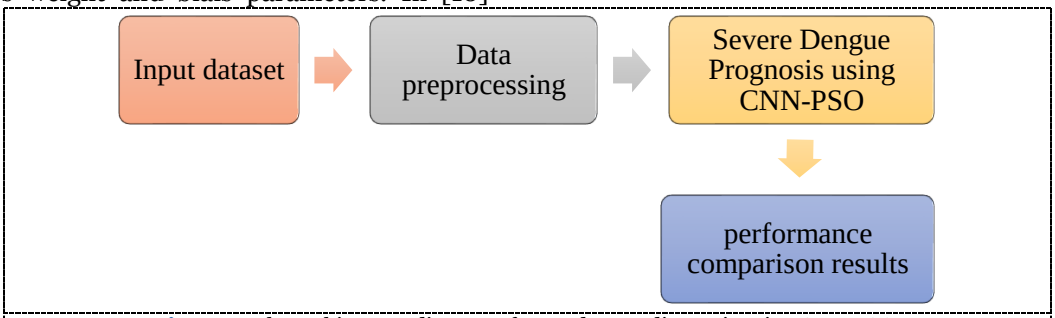

Figure 1. The architecture diagram of sever dengue diagnosis using CNN-PSO.

3.1. Patient Cohort and Preprocessing of data

The three hospitals of Recife, Brazil have been screened and invited to participate in this study for patients with dengue-related symptoms. After a detailed explanation of the study proposed, all patients who decided to participate were included in this research. FIOCRUZPE: CEP / CPQAM N011/11 and C.A.A.E. 0009.0.095.000-11, IORG0001419 reviewed and accepted the report. Whole blood was collected and tested for hematocrit and hemogram, white blood cell count, differential leukocyte count, serum albumin and serum aspartate (AST), and serum alanin transaminase (ALT) performed by standardized analytics from enrolled patients. The patient data is summarized in the patient cohort.

Data Pre-processing: The genome data is composed of 102 genotypes determined at 322 polymorphisms of the loci. The data were encrypted into indicators using a categorical schema to establish an SNP-genotype attribute that is homozygous dominant, heterozygous, or homozygous recessive. The lack of information at a forum was treated as a further category.

3.2. Dengue diagnosis Using CNN-PSO

In the original data set, 75% sample points were used to train the CNN, while the other 25% samples have been reserved for the test set. This study proposes transferring learning based on pretrained deep learning. A number of parameters must be determined before the training stage in the CNN [11] algorithms. These parameters may influence the classification results more or less depending on the application. After the random experiments, one should consider which parameters have more effect. In order to determine the best value, one could be statistically evaluated as a factor in the whole factorial test. The algorithm defaulted to set other parameters with less impact on the results.

Convolutional Neural Network: CNN is an approach to machine learning focused on the profound structure of the brain [11]. This network is essentially comprised of three layers: concentration, sub-sampling or pooling and full link layers. So each layer is briefly described in the following sections.

Convolution Layer: Convolution Layer: There are a number of n filters in each layer. The number of layers used for the convolution procedure is the

input of those filters and the depth of the feature(s) generated is proportional to the number of filters. Recognize that another filter is perceived to be a special feature in a certain data set attribute.

The output of the l-th convolution layer, represented as $Conv_j^{(l)}$, encompasses of feature attributes. It is computed as

$$Conv_i^{(l)} = bias_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} * C_j^{(l)} \quad (1)$$

Where $bias_i^{(l)}$ stands the bias matrix and $K_{i,j}^{(l-1)}$ stands convolution filter or kernel of size $a * a$ that connects the j-th feature in layer $(l-1)$ through i-th feature in the same layer also the output layer contains of feature. The input space in (1) is the primary convolutional layer, $Conv_i^{(0)} = X_i$.

The functionality is given by the kernel. Subsequently the convolution layer, the activation function will trigger the nonlinear transformation of the output of the convolutional layer:

$$Y_i^{(l)} = Y\left(Conv_i^{(l)}\right) \quad (2)$$

where $Y_i^{(l)}$ remains the output of the activation function besides $Conv_i^{(l)}$ remains the input that it receives.

Normally used activation functions are sigmoid, tanh, and rectified linear units (ReLUs). In this work, ReLUs which is signified as $Y_i^{(l)} = \max\left(0, Y_i^{(l)}\right)$ are used.

Fully Connected layer: The end layer of the CNN is a typical network of reviews with one or more hidden colum ns. Softmax activation function is used in the output layer:

$$y_i^{(l)} = \sum_{i=1}^{m_i^{(l-1)}} w_i^{(l-1)} y_i^{(l-1)} \quad (3)$$

where $w_i^{(l-1)}$ remains the weight to be modified by the completely connected sheet, the transfer function represents nonlinearity to form the representation of each class. Note that nonlinearity of the fully connected layer is not contained in its neurons, not in separate layers, as in convolutions and pooling layers. While examining for feedback signals, the development of the CNN is begun. Training takes place using the stochastic gradient descent method [11]. Then use a single example from a workout, the algorithm explores the gradients. Following planning, the parameters of CNN are calculated.

3.3. Hyper parameter tuning of CNN using PSO

The neurons, weight and bias parameters to be optimized by PSO and parameter setting of PSO are shown in Table 1.

Table 1. Parameter setting of PSO

Parameter	values
Population of PSO (p)	15
PSO stop criteria	50
Velocity coefficient (v_{c1}, v_{c2})	2
Hindering coefficient (w)	15
Max and Min value of Hindering coefficient (w_{max}&w_{min})	0.9 & 0.4 respectively
Number of particle dimension (d)	8

PSO produces and evolves several solutions to a single problem over many centuries. Each solution contains all parameters to improve the results. Weights in both layers contribute to CNN's high precision. A single PSO solution would therefore be usable in all CNN weights. The CNN has four levels (1 input, 2 hidden and 1 output), according to the network structure discussed in the previous tutorial and shown in the figure below. Recall that PSO uses a fitness function of precision to assign each solution a fitness value and the higher the fitness value, the better the response. The following steps are given:

Step 1 (Solution Defining Space and Fitness Functions): Pick optimized parameters. Put this in an appropriate range with the optimum solution. Instead, in a multidimensional optimisation, set a minimum and maximum value for each dimension.

Step 2 (Initialization phase): Determine population size and maximum iteration number. Set the starting position for each particle and start speed. Accuracy is used to calculate the error between the prediction value and the true data for every particle.

Step 3 (optimization phase): first, find the best part position and best part position in the swarm in its history. Second, positions and speeds are upgraded by equations (4) to (6) where the weight of inertia and the learning factor are concerned.

$$x_k^{s+1} = x_k^s + v_k^{s+1} \quad (4)$$

$$v_k^{s+1} = wv_k^s + v_{c1}.rand.\left(x_{pbest(k)}^s - x_k^s\right) + v_{c2}.rand.\left(x_{gbest(k)}^s - x_k^s\right) \quad (5)$$

$$w = w_{max} - (w_{max} - w_{min}) * \frac{s}{I} \quad (6)$$

In Equations (4) to (6) above, v_k^{t+1} and x_k^{t+1} are the s is the speed and position components of the kth particle, v_{c1}, v_{c2} are the Velocity coefficients and the $pbest(k)$ and $gbest(k)$ are personal best and global best particle. Here the w hindering coefficient as it helps the particles to move by hindering towards better positions and finally rand is a uniform random value between 0 and 1.

Step 4: If the evaluation function (predicted accuracy of the training information) is converging, the optimization will end; otherwise, proceed to step 3 and Step 4 is finished finally.

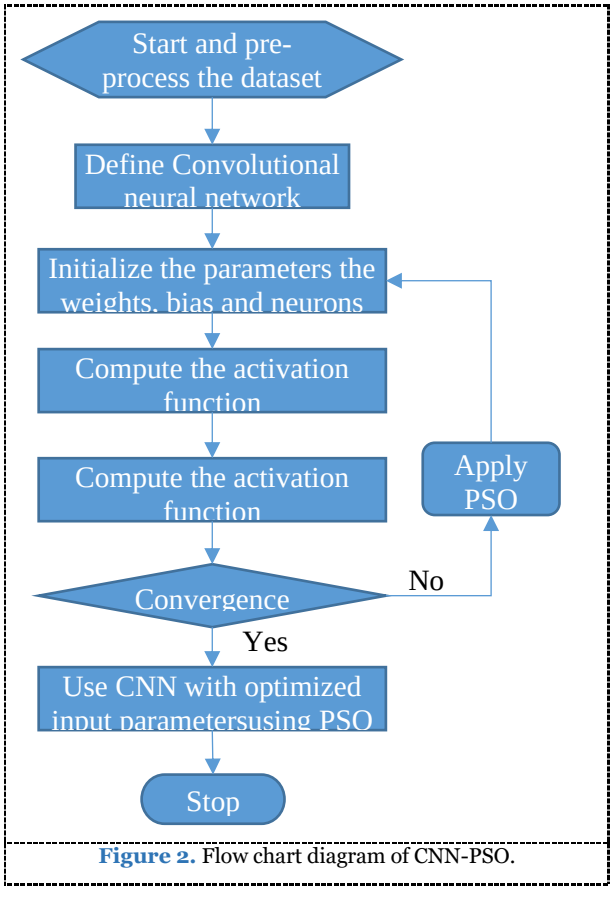

Figure 2. Flow chart diagram of CNN-PSO.

Therefore the description of each particle is such that the one part contains variables (particles) with discrete values (that represent the neuron numbers for each layer). The other part contains variables with continuous values (that represent the weights and bias values). At this stage, this specification is used to create one random initial population. Using particulate values in the determined population, the CNN architecture and training weights and distortions are specified. PSO trains its fitness function for every chromosome, which is determined for this population, then via the standard input data, heights, weights and biases proposed for this. Figure 2 shows the flow chart of CNN-PSO.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed CNN-PSO is evaluated in this section, and the performance results are compared with existing MPNN [6], ANFIS [8] and PSO-ANN [9] schemes. The performance measurement is done in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (7):

$$Precision = \frac{TP}{FP+TP} \quad (7)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (8):

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

F-measure: this is also referred to as F 1-score, and as in equation (9) is the harmonic mean of precision and recall:

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (9)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (10):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (10)$$

Where true positive (TP) samples are properly classified as no dengue, false positive (FP) samples are incorrectly classified as dengue, True negative (TN) samples are properly classified as dengue, and false negatives (FN) are incorrectly classified as dengue.

4.1. Precision Rate comparison

From the above Figure 3, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as MPNN, ANFIS, PSO-ANN and CNN-PSO. When the number of records increases according to the precision value. From this graph, it is learned that the proposed CNN-PSO offers 94%

higher precision than previous methods that yield better results in the classification of lung nodules due to hyper parameter optimization of CNN using PSO. The numerical results of Precision Comparison is shown in Table 2.

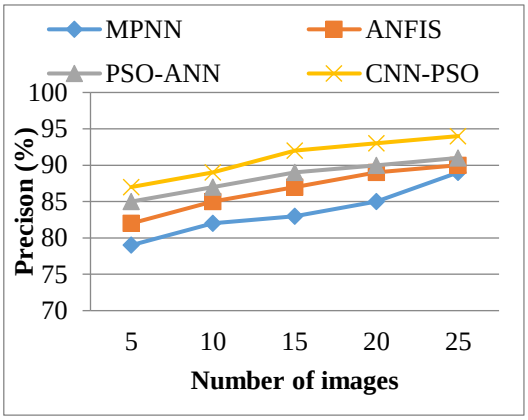

Figure 3. Representation of Precision Comparison

Table 2. The numerical results of Precision Comparison

No. of images	MPNN	ANFIS	PSO-ANN	CNN-PSO
5	79	82	85	87
10	82	85	87	89
15	83	87	89	92
20	85	89	90	93
25	89	90	91	94

4.2. Recall comparison

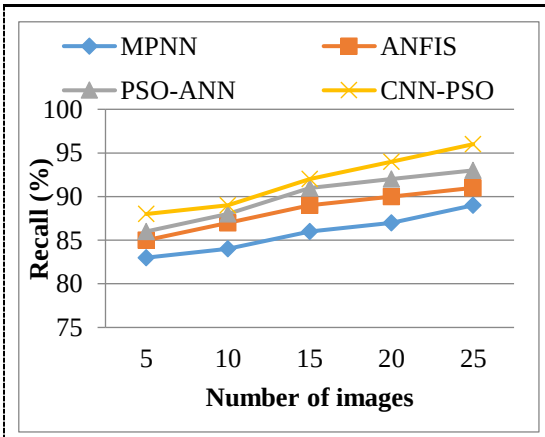

Figure 4. Representation of Recall Comparison

From the above Figure 4, the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as MPNN, ANFIS, PSO-ANN and CNN-PSO. Increasing the number of photographs often increases the correct value for the recall. Through this graph, it is discovered that the current CNN-PSO offers recall 96% higher than previous methods. The explanation for this is that the CNN-PSO extracts the features directly which will enhance the detection of severe dengue. The numerical results of Recall Comparison is shown in Table 3.

Table 3. The numerical results of Recall Comparison

No. of images	MPNN	ANFIS	PSO-ANN	CNN-PSO
5	83	85	86	88
10	84	87	88	89
15	86	89	91	92
20	87	90	92	94
25	89	91	93	96

4.3. F-measure Rate comparison

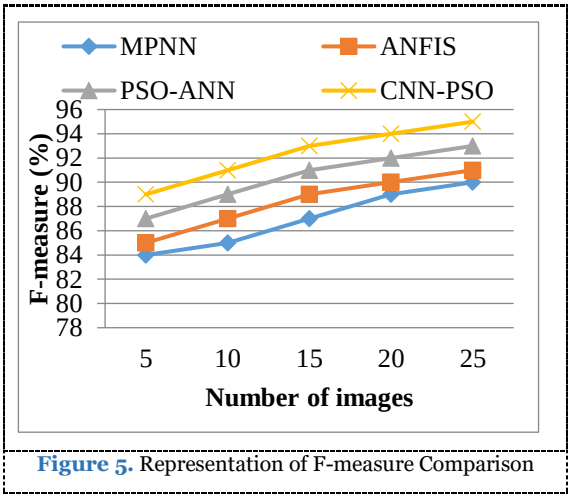

Figure 5. Representation of F-measure Comparison

From the above Figure 5, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as MPNN, ANFIS, PSO-ANN and CNN-PSO. When the number of data is increased, and the f-measure value is increased accordingly.

Table 4. The numerical results of F-measure Comparison

No. of images	MPNN	ANFIS	PSO-ANN	CNN-PSO
5	84	85	87	89
10	85	87	89	91
15	87	89	91	93
20	89	90	92	94
25	90	91	93	95

From this graph it is learned that the proposed CNN-PSO offers 95% higher f-measurement than previous methods. Therefore the proposed CNN-PSO algorithm is stronger than the current algorithms in terms of better performance of prognosis of severe dengue. The numerical results of F-measure Comparison is shown in Table 4.

4.4. Accuracy comparison

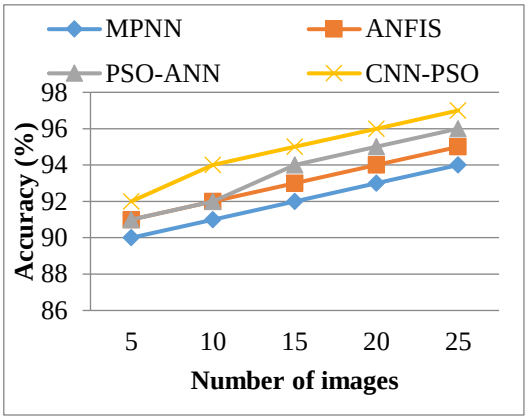

Figure 6. Representation of Accuracy Comparison

From the above Figure 6, the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as MPNN, ANFIS, PSO-ANN and CNN-PSO. From this graph it is known that the proposed CNN-PSO algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better template matching results. This is due to the automatic extraction of the function in the CNN-PSO algorithm, which increases the severe dengue prognosis. The numerical results of Accuracy Comparison is shown in Table 5.

Table 5. The numerical results of Accuracy Comparison

No. of images	MPNN	ANFIS	PSO-ANN	CNN-PSO
5	90	91	91	92
10	91	92	92	94
15	92	93	94	95
20	93	94	95	96
25	94	95	96	97

V. CONCLUSION AND FUTURE WORK

In this work, a forecast of severe dengue, based on CNN-PSO, is proposed. The primary advantages of a CNN-based genome method for the prediction of serious dengue development are that it can also be implemented before infection at any stage of disease, and that human samples can be widely selected. The results show that the presented methodology offers a robust tool for pronouncing dengue severity. Experimental data show that the built CNN-PSO can detect most of the nodules with a high 97% accuracy rate. This success is mainly due to the deep CNN structure, which uses the capacity to extract different level features and to the use of PSO to modify hyperparameters that has resulted in better generalization. Smart techniques along with evolutionary algorithms can be used for faster calculation by choosing the optimisation required in different aspects. Considering multi-target distinguishing changes, the measurement accuracy is a good attempt.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank their present employer for providing support while carrying out this research work.

REFERENCES

[1]. Ta-Chien Chan, Tsuey-Hwa Hu, Jing-Shiang Hwang, "Daily forecast of dengue fever incidents for urban villages in a city", International *Journal of Health Geographics*, 2015.

[2]. Khan MI Anwar E Agha A et al. Factors predicting severe dengue in patients with dengue fever. Mediterr J Hematol Infect Dis 2013; 5:e2013014.

[3]. Ho TS Wang SM Lin YS Liu CC . Clinical and laboratory predictive markers for acute dengue infection. *J Biomed Sci,*2013; 20:75.

[4]. Falconar AK Romero-Vivas CM . Simple prognostic criteria can definitively identify patients who develop severe versus non-severe dengue disease, or have other febrile illnesses. *J Clin Med Res*, 2012; 4:33–44.

[5]. Faisal T, Taib MN, Ibrahim F. Adaptive Neuro-Fuzzy Inference System for diagnosis risk in dengue patients. *Expert Systems with Applications*. 2012 Mar 1;39(4):4483-95.

[6]. Faisal T, Taib MN, Ibrahim F. Neural network diagnostic system for dengue patients risk classification. *Journal of medical systems*. 2012 Apr 1;36(2):661-76.

[7]. Jiji GW, Lakshmi VS, Lakshmi KV, Priya SS. Diagnosis and Prognosis of the Arbovirus-Dengue using Intelligent Algorithm. *Journal of The Institution of Engineers (India):* Series B. 2016 Jun 1;97(2):115-20.

[8]. Saikia D, Dutta JC. Adaptive Network Based Fuzzy Inference System for Early Diagnosis of Dengue Disease. *InAdvances in Computer and Computational Sciences* 2017 (pp. 721-728). Springer, Singapore.

[9]. Gambhir S, Malik SK, Kumar Y. PSO-ANN based diagnostic model for the early detection of dengue disease. *New Horizons in Translational Medicine*. 2017 Nov 1;4(1-4):1-8.

[10]. Rao VS, Kumar MN. A new intelligence-based approach for computer-aided diagnosis of dengue fever. *IEEE transactions on information technology in biomedicine*. 2011 Oct 17;16(1):112-8.

[11]. Murphy J. An overview of convolutional neural network architectures for deep learning. *Microway Inc*. 2016.

Cite this article as: Mustafa M. et al. Hybrid Convolutional Neural Network with PSO Based Severe Dengue Prognosis Method in Human Genome Data. J. Comput. Sci. Intell. Technol. 2020; 1(1): 22-28. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 1, pages: 29 – 36, March 2020

A New Framework for Anomaly Detection in NSL-KDD Dataset using Hybrid Neuro-Weighted Genetic Algorithm

[1]Muneeshwari P, & [2]Kishanthini M

[1]Department of Information Technology, PSNA College of Engineering and Technology, Dindigul, Tamilnadu, India.

[2]Department of Computer Science and Engineering, Amrita College of Engineering and Technology, Nagercoil, Tamilnadu, India.

Corresponding Author: radhamunishapcse@gmail.com

Received: *05.01.2020,*
Revised: *08.02.2020,*
Accepted: *17.03.2020,*
Published: *30.03.2020*

DOI:
10.53409/mnaa.jcsit1105

Abstract: There are an increasing number of security threats to the Internet and computer networks. For new kinds of attacks constantly emerging, a major challenge is the development of versatile and innovative security-oriented approaches. Anomaly-based network intrusion detection techniques are in this sense a valuable tool for defending target devices and networks from malicious activities. With testing dataset, this work was able to use the NSL-KDD data collection, the binary and multiclass problems. With that inspiration, data mining techniques are used to offer an automated platform for network attack detection. The system is based on the Hybrid Genetic Neuro-Weighted Algorithm (HNWGA).In this weighted genetic algorithm is used for the selection of features and in this work a neuro-genetic fuzzy classification algorithm has been proposed which is used to identify malicious users by classifying user behaviors. The main benefit of this proposed framework is that it reduces the attacks by highly accurate detection of intruders and minimizes false positives. The evaluation of the performance is performed in NSL-KDD dataset. The experimental result shows of that the proposed work attains better accuracy when compared to previous methods. Such type of IDS systems are used in the identification and response to malicious traffic / activities to improve extremely accuracy.

Keywords: Data mining, Hybrid Genetic Neuro-Weighted Algorithm, neuro-genetic fuzzy classification and NSL-KDD dataset.

I. INTRODUCTION

Network intrusion detection (IDS), by detecting malicious users, are useful for providing protection to allow only legitimate users and detach malicious users further. The use of IDSs[1] effectively meets safety criteria such as confidentiality, non-repudiation and authentication. Both servers or on network nodes, IDSs can be installed. Innocence and external threats are marked. Due to their membership in businesses and associations, network users misuse their privileges and certificates issued by internally targeted individuals.

This form of malicious intrusion would allow network resources to be leveraged through network services to be disrupted. The external user should then define attacks on the basis of external user anomalies. In this case, network usage habits for these attackers are monitored for a defined length, such that attackers can be differentiated from legitimate users. The literature already contains a wide range of methods and tools for intrusion detection. Until then, most current tools rely on the analysis of a benchmark dataset on specified types of attacks and do not conduct smart, soft calculation-based analysis[2]. New techniques must therefore be put forward, that could analyze all types of attacks effectively by applying intelligent soft-computing techniques.

However, numerous researchers utilize methods for selection of features to increase classification algorithms and IDS[3] performance. With certain cases, knowledge collection or information processing is used to make final decisions. In addition, the values in the dataset come from various attributes and their data types are also not standardized. In order to handle this issue , it is important for each attribute to be normalized and emphasized based upon its importance in the classification process.

This research successfully recognize intrusion into networks through a new neuro-genetic hybrid architecture called HNWGA. The proposed HNWGA includes various components and, above all, feature selection and assault classification. Such components have been developed using smart approaches and thus a new genetic algorithm (GFCA) and a smart classification algorithm (GFCA) have been proposed in the present paper to improve detection accuracy. These algorithms include a new genetic algorithm based upon weighted genetic algorithms (WGA). The proposed WGA is here used to define the optimum number of features which are ideal for classifying intruders successfully. The rest of the paper is structured as follows: In section 2 the related work of IDS is discussed. The proposed IDS detection mechanism using HNGA is described in section 3. The experimental results and discussion is discussed in section 4. The conclusion and future work is given in section 5.

II. RELATED WORK

In literature there are a few works aimed at detecting network intrusion. In [4], an efficient, reliable classifier is developed to evaluate a visit to a network as normal and not to the ant-colony algorithm and the support vector machine (SVM). In [5] a technical approach used by PCA to choose the sub-set of SVM, the classification feature and the selection of feature subsets based on their own values was proposed. [6], proposed algorithms such as Efficient Data Adapted Decision Tree(EDADT), Hybrid IDS, Semi-supervised methods and Hopping Period Alignment and Adjustment (HOPERAA) varying algorithms respectively. In [7], the new hybrid intrusion detection method is proposed, consisting of a C4.5 decision tree algorithm and a decomposition structure anomaly model, integrated hierarchically. Next, for desiccated subsets, SVM models are created. Two new methods for hybrid intrusive detection are reported in the study presented in [8]: one is based on gravitational search, and one on a mix of GS and GSPSO. In [9], suggested a deep learning approach on recurrent neural networks (RNN-IDS) for intrusion detection. Still be extremely cautious to reduce the training time with the use of GPU acceleration, prevent explosions and disappearances of gradients and study LSTM's classification performance in the field of the bidirectional RNN algorithm. In [10], a hybrid intrusion detection model of several levels, which uses a vector and extreme learning machine to enhance the effectiveness of the detection of external and internal attacks. Despite a large number of existing works, only few types of attacks and other kinds of attacks were taken into consideration by most of the existing systems. In this work , a new IDS is therefore proposed that considers all kinds of known dataset attacks as well as feature selection to enhance the precision of classification.

III. PROPOSED METHODOLOGY

HNWGA-based IDS consists of two primary components, the feature-selection framework for the optimum number of features and the data classification framework using the selected functions. Currently, the necessary data is obtained from the NSL-KDD dataset. The proposed algorithm of WGA function selection will be included in this context in order to choose the appropriate number of features from 41, and the proposed HNWGA algorithm can then be used to classify the data set effectively and the design diagram is shown in Figure 1. In the final judgment of the network data received from IDs by the fluctuating rules manager and the dynamic rule base, this Judgment Manager assists.

3.1. Input NSL-KDD dataset and preprocessing

Various statistical analysis showed the inherent drawbacks in the KDD cup 99 dataset which affect the accuracy of many researchers' IDS detection systems. Data set NSLKDD is its predecessor's refined version. The complete KDD dataset contains an essential record. A collection of files for the researchers can be downloaded. In [11], the details of the attributes are indicated: the name, description of the attributes and sample data. Pre-processing data set: The initial pre-processing steps are as follows:

- Dataset upload: The planned data set is uploaded for use in the data mining process in this step.
- Features for extraction: the desired features are chosen from a data set up in step 1. A special feature, a subset of properties or all features may be included in the extraction basis.
- Feature roles determination: The feature roles are indicated in this stage. "Roles" identify a specific feature identification number and determine whether it is regular, special, labeled, etc.

- Conversion of nominal data to numerical: nominal data must be converted to numerical data.
- Standardization: Feature values based on Z-transformation are normalized in this step.

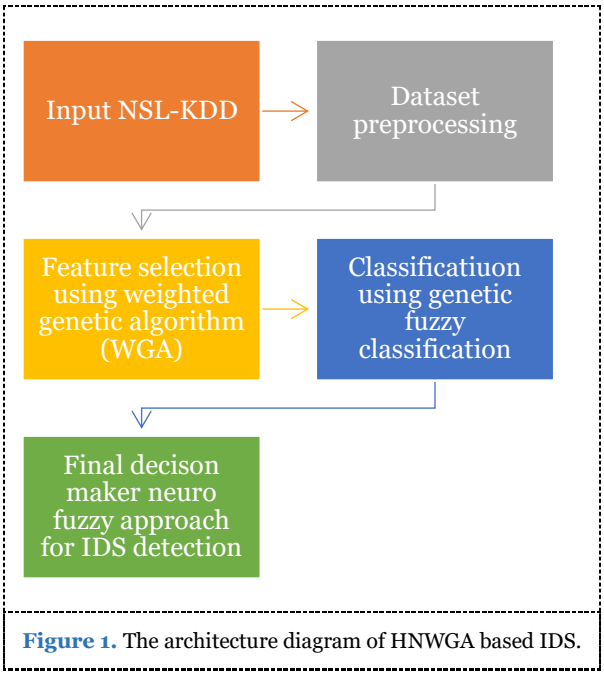

Figure 1. The architecture diagram of HNWGA based IDS.

3.2. Feature Selection Using Weighted Genetic Algorithm

To The input data is converted into binary formats and used for generation formation in the genetic algorithm process. In this model, the proposed algorithm for each iteration is taken from two chromosomes pertaining to two datasets, and they are assessed for physical accuracy, while the algorithm selects two parents when accommodate. If one of them is not suitable, the following dataset record will be considered for the next parent. This process is repeated with the application of fitness values, the choice of two parents, cross-sectional operations and mutations before fitness is assessed. The function set is made up of all attributes chosen by the weighted genetic algorithm. The initial population was created through transforming the data values to binary values. The number of 1s is created for each individual based on their original values, for various characteristics in sub-sets. The suggested weighted average fitness assessment function has been used to determine the weighted average accuracy and number one. Variations of the current generations of different chromosome groups are analyzed using the fusion and mutation operators. Crossover is performed in the latter part of the chromosome by means of uniform crossover operation and mutation. The selection is done by selecting a tournament in which the algorithm selects chromosome subsets for the whole population. The working principle flow diagram of WGA is shown in Figure 1. Table 1 provides pseudo code for the selection of features using weighted genetic algorithms. The initial population was created through transforming the data values to binary values. The number of 1s is created for each individual based on their original values, for various characteristics in sub-sets. The suggested weighted average fitness assessment function has been used to determine the weighted average accuracy and number one. Variations of the current generations of different chromosome groups are analyzed using the fusion and mutation operators. Crossover is performed in the latter part of the chromosome by means of uniform crossover operation and mutation. The selection is done by selecting a tournament in which the algorithm selects chromosome subsets for the whole population. The working principle flow diagram of WGA is shown in Figure 1. Table 1 provides pseudo code

for the selection of features using weighted genetic algorithms.

Table 1. The pseudo code of feature selection using weighted genetic algorithms

Input: NSL-KDD features (F) (41 features), maximum number of iterations (max_generations), number of records (population size taken from the dataset), crossover probability ($\mathcal{CP}$), mutation probability ($\mathcal{MP}$).

Output: Optimal number of selected features

1. Initialize the chromosome population consisting of 41 attributes
2. Convert each value of the attributes into binary so that chromosomes can be either '0' or '1.'
3. Initialise the weights $w_1 = 0.6$ and $w_2 = 0.4$ to the chromosomes (C_i)
4. Do for every single chromosome (C_i)
5. Calculate uniform crossover on C_i with a $\mathcal{CP}$ probability.
6. Calculate mutation operator with a probability of $\mathcal{MP}$ to the last bits of the C_i.
7. Evaluate the weighted average fitness evaluation function $\mathcal{F}(i) = \frac{\left[w_1 * accuracy(i) + w_2 * \left(\frac{1}{number\ of\ ones}\right)\right]}{(w_1 + w_2)}$
8. Condition check $\mathcal{F}(i) > threshold$ then $\mathcal{F}(i) = \mathcal{F}(i)\ Feature\ set$
9. Using tournament selection from Feature set and choose the top best chromosomes in the new population as the optimal features.
10. Repeat the steps from 3 to 9 until the stop criterion is met

Produce the results of an optimal selection of features.

The input data is converted into binary formats and used for generation formation in the genetic algorithm process. In this model, the proposed algorithm for each iteration is taken from two chromosomes pertaining to two datasets, and they are assessed for physical accuracy, while the algorithm selects two parents when accommodate. If one of them is not suitable, the following dataset record will be considered for the next parent. This process is repeated with the application of fitness values, the choice of two parents, cross-sectional operations and mutations before fitness is assessed. The function set is made up of all attributes chosen by the weighted genetic algorithm.

The initial population was created through transforming the data values to binary values. The number of 1s is created for each individual based on their original values, for various characteristics in sub-sets. The suggested weighted average fitness assessment function has been used to determine the weighted

average accuracy and number one. Variations of the current generations of different chromosome groups are analyzed using the fusion and mutation operators. Crossover is performed in the latter part of the chromosome by means of uniform crossover operation and mutation. The selection is done by selecting a tournament in which the algorithm selects chromosome subsets for the whole population. The working principle flow diagram of WGA is shown in Figure 1. Table 1 provides pseudo code for the selection of features using weighted genetic algorithms.

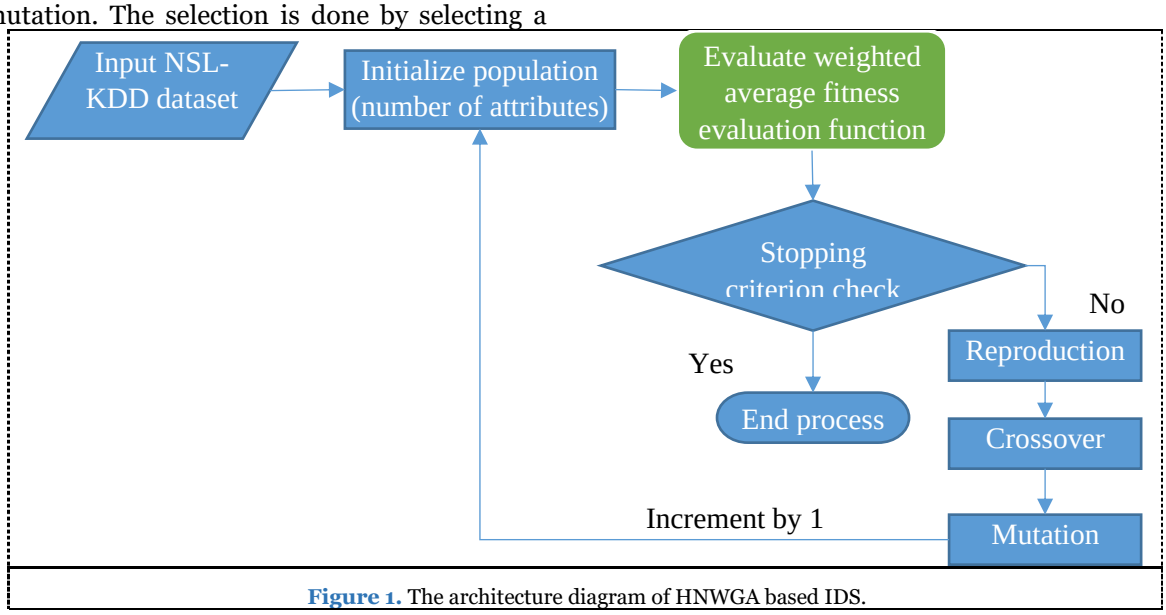

Figure 1. The architecture diagram of HNWGA based IDS.

3.3. Hybrid Neuro-weighted genetic fuzzy classification for IDS detection

In this work a new algorithm with fuzzyrules was proposed and evaluated with the benchmark data collection, called the neuroweighted genetic fuzzy classification algorithm (HNWGA). In this classification algorithm, one input layer, one output layer, and two hidden layers were used for back propagation neural networks (BPNNs). As something of an activation function for neural network modeling, the exponential function was being used. Weight modification is often done by means of genetic algorithms with fuzzy rules and the fuzzy rules are used for ultimate choice. The proposed IDS algorithm, HNWGA, is shown in Table 2.

Table 2. The pseudo code of feature selection using weighted genetic algorithms

Input: NSL-KDD with records $\boldsymbol{R_i}, \boldsymbol{i} = \boldsymbol{1, 2 \ldots, n}$, with attributes $\boldsymbol{A_j}, \boldsymbol{j} = \boldsymbol{1, 2 \ldots, m}$, Optimal number of selected features, weights for chromosome $\boldsymbol{x}$ are $\boldsymbol{w_1} = \boldsymbol{0.6}$ and $\boldsymbol{w_2} = \boldsymbol{0.4}$.
Output: The anomaly detection results with attack types
From the KDD cup dataset select some amount of records from the total records at random.
Train data set utilizing neural networks with optimum features in back propagation
Initialize population size, binary form attributes, probability of crossover and probability of mutation.
Read genetic algorithm fitness function as $\mathbb{F}(\boldsymbol{x}) = \frac{(\boldsymbol{w_1} * \boldsymbol{no.of\ zeros}) + (\boldsymbol{w_2} * \boldsymbol{no.of\ ones})}{(\boldsymbol{w_1} + \boldsymbol{w_2})}$
For $\boldsymbol{i} = \boldsymbol{1\ to\ n}$ do
Start substituting the initial BPNN class labels with the training that was conducted by means of first labeling.
By applying union operation enhance new data samples into the training data set.
Repeat steps 14 and15 until the stopiing criterion is obtained
Generate fuzzy rules by applying trapezoidal membership to Training data [12].
Apply fuzzy rules to make weight adjustment decisions
Apply activation function to hidden layers output
Perform the process of defuzzification.
Load the attack types which are classified.
Form rules and store them in fuzzy rule base for testing.

Read rules from the base of Fuzzy rules and apply to test data.

Presentation of final result, showing the IDS with its attack types.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In NSL-KDD 1999 cup data set was used in this work to evaluate the feature selection algorithm and classification algorithm developed for the development of an IDS containing five classes (probe, U2R, R2L, DoS, and normal). In this section, the proposed HNWGA quality is assessed using traditional methods such as SVM-IDS [4], GSPSO-IDS [8], and RNN-IDS [9] with certain parameters such as accuracy, f-measure, accuracy and recall.

Precision: It represents the proportion of positive samples correctly classified to the total number of positive samples predicted as shown in equation (1):

$$Precision = \frac{TP}{FP+TP} \quad (1)$$

Recall: a classifier's recall represents the positive samples correctly classified to the total number of positive samples and is estimated as in equation (2):

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

F-measure: this is also referred to as F 1-score, and is the harmonic mean of precision and recall as in equation (3):

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (3)$$

Accuracy: This is one of the most commonly used classification performance measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (4):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Where true positive (TP) samples are properly classified as normal, false positive (FP) samples are incorrectly classified as abnormal, true negative (TN) samples are properly classified as abnormal, and false negatives (FN) are incorrectly classified as normal.

4.1. Precision Rate comparison

From the above Figure 2, the graph explains the comparison of precision for the number of records in the specified datasets. Such methods as SVM-IDS, GSPSO-IDS, RNN-IDS and HNWGA are executed. When it increases the number of records according to the precision value. From this graph, it is learned that, due to optimal feature selection technique.

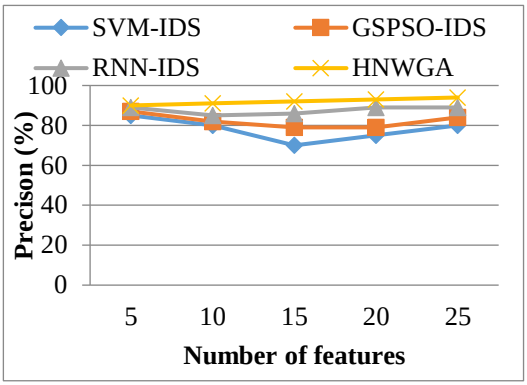

Figure 2. Representation of Precision Comparison

The proposed HNWGA provides 94% higher precision than the previous methods that produce better results in attack detection. The numerical results of Precision Comparison is shown in Table 3.

Table 3. The numerical results of Precision Comparison

No.of features	SVM-IDS	GSPSO-IDS	RNN-IDS	HNWGA
5	85	87	89	90
10	80	82	85	91
15	70	79	86	92
20	75	79	89	93
25	80	84	89	94

4.2. Recall comparison

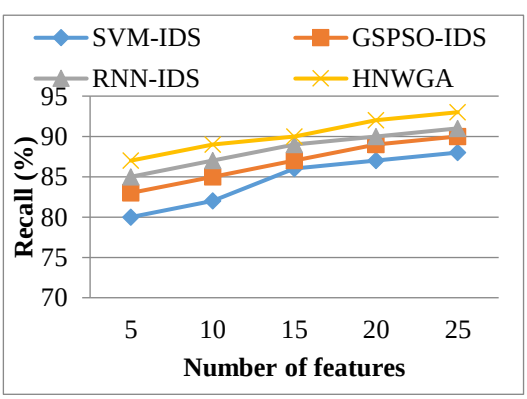

Figure 3. Representation of Recall Comparison

The graph explains from the above Figure 3 that the recall comparison for the number of records in the specified datasets. Such methods as SVM-IDS, GSPSO-IDS, RNN-IDS and HNWGA are executed. Increasing the number of images also increases the corresponding recall value. It is learned from this graph that the proposed HNWGA provides 93% higher recall than previous methods. The reason for this is that the WGA produces the optimal features that will improve the results of attack detection. The numerical results of Recall Comparison is shown in Table 4.

Table 4. The numerical results of Recall Comparison

No.of features	SVM-IDS	GSPSO-IDS	RNN-IDS	HNWGA
5	80	83	85	87
10	82	85	87	89
15	86	87	89	90
20	87	89	90	92
25	88	90	91	93

4.3. F-measure Rate comparison

From the above Figure 4, the graph explains the comparison of the f-measure for the number of images in specified datasets. Such methods as SVM-IDS, GSPSO-IDS, RNN-IDS and HNWGA are executed. When the number of data is increased and correspondingly the f-measure value is raised. It is learned from this graph that the proposed HNWGA provides 96 per cent higher f-measurement than previous methods.

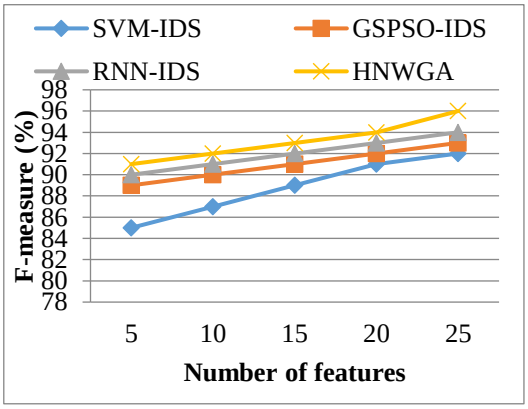

Figure 4. Representation of F-measure Comparison

Thus the proposed HNWGA algorithm is greater in terms of better results of attack detection than the existing algorithms. The reason Neuro fuzzy's parameter is optimized with genetic algorithm which will enhance the results of attack detection. The numerical results of F-measure Comparison is shown in Table 5.

Table 5. The numerical results of F-measure Comparison

No.of features	SVM-IDS	GSPSO-IDS	RNN-IDS	HNWGA
5	85	89	90	91
10	87	90	91	92
15	89	91	92	93
20	91	92	93	94
25	92	93	94	96

4.4. Accuracy comparison

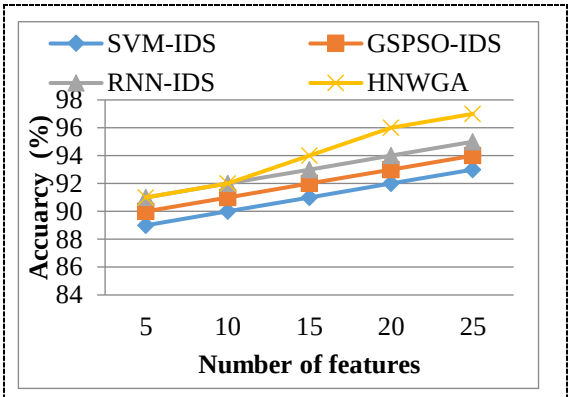

Figure 5. Representation of Accuracy Comparison

From the above Figure 5, the graph explains the comparison of processing time for the number of images in the specified datasets. Such methods as SVM-IDS, GSPSO-IDS, RNN-IDS and HNWGA are executed. It is learned from this graph that the proposed HNWGA algorithm is higher than the existing algorithms in terms of better template matching results with a high precision rate of 97%. The reason is that existing approaches also have a low success rate which has a high likelihood of causing misdetection of emerging changes. This is due to the use of the most important features selected by the feature selection algorithm which uses intelligent agents for decision making and, in addition, the use of fuzzy rules and genetic algorithm in the classification algorithm increases the classification accuracy resulting in an increase in in intrusion detection accuracy. The numerical results of Accuracy Comparison is shown in Table 5.

Table 5. The numerical results of Accuarcy Comparison

No.of features	SVM-IDS	GSPSO-IDS	RNN-IDS	HNWGA
5	89	90	91	91
10	90	91	92	92
15	91	92	93	94
20	92	93	94	96
25	93	94	95	97

V. CONCLUSION AND FUTURE WORK

A new HNWGA framework for an IDS was suggested in this work. Towards this end , a new algorithm called WGA has been suggested for the selection of features which will enhance detection accuracy, network efficiency, and optimal selection of features. Furthermore, there has been a special proposal on a new classification algorithm called GFCA, that also aims to improve the accuracy of intrusion detection. Experiments carried out in this paper show that the GFCA increases the accuracy of classification and that the classification time when

selected features are used. Ultimately, the work proposed named HNWGA was tested and its performance analyzed by means of a precise analysis and comparable with SVM-IDS, GSPSO-IDS, RNN-IDS in constant environments. Experimental studies in this study have shown that the proposed algorithm of classifying results is more accurate than the other three classifiers. The key benefit of the proposed IDS system is that the identification is more precise, false positive and deduction times are of. Future research can be carried out in this way to help handle uncertainty by using the adaptive fuzzy inference model.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank their present employer for providing support while carrying out this research work.

REFERENCES

[1]. Ghorbani AA, Lu W, Tavallaee M. Network intrusion detection and prevention: concepts and techniques. Springer Science & Business Media; 2009 Oct 10.

[2]. Liao HJ, Lin CH, Lin YC, Tung KY. Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications. 2013 Jan 1;36(1):16-24.

[3]. Ganapathy S, Kulothungan K, Muthurajkumar S, Vijayalakshmi M, Yogesh P, Kannan A. Intelligent feature selection and classification techniques for intrusion detection in networks: a survey. EURASIP Journal on Wireless Communications and Networking. 2013 Dec 1;2013(1):271.

[4]. Li Y, Xia J, Zhang S, Yan J, Ai X, Dai K. An efficient intrusion detection system based on support vector machines and gradually feature removal method. Expert Systems with Applications. 2012 Jan 1;39(1):424-30.

[5]. Ahmad I, Hussain M, Alghamdi A, Alelaiwi A. Enhancing SVM performance in intrusion detection using optimal feature subset selection based on genetic principal components. Neural computing and applications. 2014 Jun 1;24(7-8):1671-82.

[6]. Nadiammai GV, Hemalatha MJ. Effective approach toward Intrusion Detection System using data mining techniques. Egyptian Informatics Journal. 2014 Mar 1;15(1):37-50.

[7]. Kim G, Lee S, Kim S. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications. 2014 Mar 1;41(4):1690-700.

[8]. Dash T. A study on intrusion detection using neural networks trained with evolutionary algorithms. Soft Computing. 2017 May 1;21(10):2687-700.

[9]. Yin C, Zhu Y, Fei J, He X. A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access. 2017 Oct 12;5:21954-61.

[10]. Al-Yaseen WL, Othman ZA, Nazri MZ. Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system. Expert Systems with Applications. 2017 Jan 1;67:296-303.

[11]. L. Dhanabal and S. P. Shantharajah, "A Study on NSLKDD Dataset for Intrusion Detection System Based on Classification Algorithms," vol. 4, no. 6, pp. 446–452, 2015.

[12]. Palanivel K. Fuzzy commercial traveler problem of trapezoidal membership functions within the sort of α optimum solution using ranking technique. Afrika Matematika. 2016 Mar 1;27(1-2):263-77.

Cite this article as: Muneeshwari P, Kishanthini M. A New Framework for Anomaly Detection in NSL-KDD Dataset using Hybrid Neuro-Weighted Genetic Algorithm. J. Comput. Sci. Intell. Technol. 2020; 1(1): 29-36. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 2, pages: 01–08, September 2020

A Survey on Security Risks in Internet of Things (IoT) Environment

Mugesh Ravi

Faculty of Informatics, Eötvös Loránd University, Budapest, Hungary.

***Corresponding Author: mugesh.ravi@outlook.com*

Received: *27.06.2020,*
Revised: *20.07.2020,*
Accepted: *17.08.2020,*
Published: *31.09.2020*

DOI:
10.53409/mnaa.jcsit20201201

Abstract: This analysis reviews the management of vulnerabilities and security risks of Internet of Things (IoT). This paper provides an overview, which it reveals the recent Internet's growth and how it has transformed our lives in various, unforeseen dimensions and how it has given rise to IoT. The introduction part focuses on providing an analysis on literature by presenting a short IoT history, some technical information on security protocols, and IoT hardware problems. The section on survey is where similar literatures on specific concepts are reviewed by describing the vulnerabilities and threats of IoT systems, and then reviewed risk management mechanisms for both information technologies and information protection. After the review, the analysis and discussion segment addressed and evaluated the details contained in the literature review. In this paper, a new risk management strategy uniquely designed for each IoT system is proposed. Then proposed work is evaluated by discussing the advantages and concluded the analysis and the future work.

Keywords: IoT, Security, Threats, Vulnerability, Risk Management Framework

I. INTRODUCTION

The Internet is undoubtedly one of the greatest innovations of mankind. It has given us several benefits, which are nothing but a fantasy recently. It's difficult on its own to deliver a message to someone across the globe, so doing it in a few milliseconds was undoubtedly a wonder. Such technological innovation has, understandably, altered many facets of our lives. Newsletters, Radio stations, Cable TV, and postal mails are all been part of the history as they are alternated by podcasts, internet news, emails, and streaming services. Most of our daily tasks include utilizing the Internet in anyway, without even noticing it. The list doesn't even end here. Small, inexpensive, and often powerful appliances are built into our watches, TVs, refrigerators, and also toasters. They can be very easy, due to their wireless existence.

It is understood, however, that convenience is typically exchanged for protection. These devices also have very small processing and memory capacities and are regularly referred to as devices or IoT of the Internet of Things. It was estimated in a report by Cisco Inc. that the number of Internet-connected devices exceeded the people population back in 2008 and will cross more than fifty billion by 2020. This highlights the value of controlling these devices' security threats and vulnerabilities, as they are an incorporated segment of our daily lives. The issue with these systems was that frequently have very constrained capacities for storage and processing. This implies that, in most instances, they have limited tools necessary to create safe communication. This makes them unprotected to different forms of attacks, potentially placing the consumer at privacy loss risk. One might discuss that it was not just a losing our privacy term that we should think about if an IoT system is compromised. Since IoT devices are often used in very sensitive settings, such as healthcare, it may exactly be a subject of life or death to properly protect them.

For these and several other purposes, the significance of providing a tailored system for handling IoT's security vulnerabilities and risks is strongly emphasized.

IoT is not a modern idea. In reality, John Romkey presented the first IoT system at the

INTEROP conference in 1989. Like a toaster that might be remotely switched on or off using the Internet. It has been linked to the computer by a stack of TCP/IP network. This occurrence set off the development of IoT. Almost a decade later, LG unveiled the first Internet-linked fridge. Although the progress of the IoT domain appeared slow, it was the major milestones after the International Telecommunications Reunion (ITR) published the study on the subject in 2005. Thereafter, the IoT field has been progressing rapidly, and in late 2008, Cisco revealed the emergence of IoT when the devices or "things" was larger than the population [3].

Today, IoT was used in a different range of ways and is very common than ever before. Though, due to the design of the devices that require efficiency of power and a factor of compact form, IoT was typically constrained about for storage and processing capacities. This has desolate consequences concerning security because secure communication algorithm typically require more computational power to operate efficiently in right time. i.e., protocol such as SSL was a key way to secure the link, primarily by encryption. SSL use the RSA cryptosystem for providing the cryptographically safe link among both ends. RSA security depends on providing huge prime numbers for measuring both private and public keys. Those big keys use a lot of memory. Moreover, SSL additionally requires output and input buffers, which again absorb memory. The ATmega328 microprocessor has only 2k bytes of SRAM are simply not provided with sufficient memory for handling SSL, or TLS. This was just one of the threats and limitations related to IoT that should be handled to achieve the objectives of the organizations.

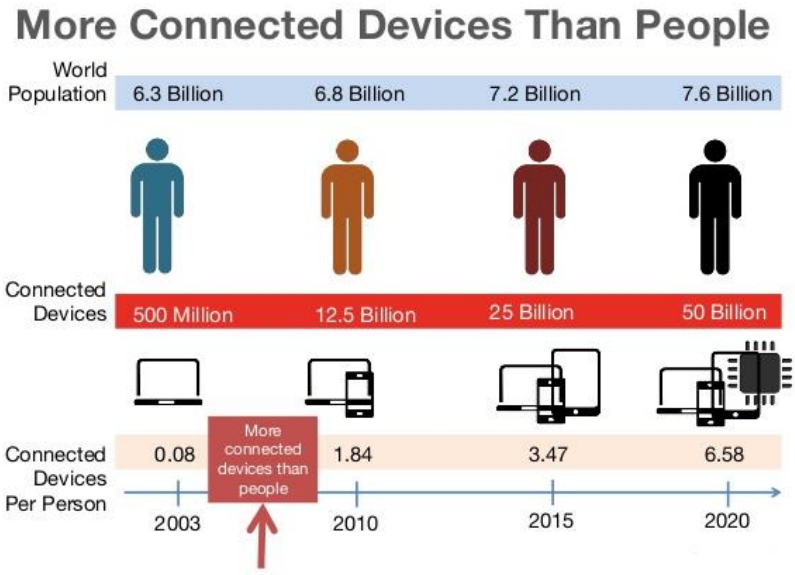

Figure.1. Growth of IoT

II. RELATED WORK

2.1. *IoT Challenges*

The management of the IoT threats and vulnerabilities was a hot field of research recently, the security issues are classified into four categories [4]:

Security issues in the application layer: involves security strategies such as trust establishment, resource exhaustion, etc.

Security issues in the architecture: often subject to application and domain scenarios.

Security issues in the communication: Responsible for the transmission of data within IoT devices or systems.

Security issues in the data protection: this was the weakest factor because the confidentiality of the data must be enclosed.

2.2. *Transcending the IoT Threats*

In order to resolve the threats, some literatures proposed the security architecture for securing the data flow of the smart grid in the home area network [5, 15], and the suggested architecture can effectively handle the transmission on the home area network

through the non-confidential and confidential principles without damaging functionality of general home area. While additionally disuse the issues of IoT and some of the disadvantages, several suggestions came across, like [6]. IoT devices must have security development and all that relates to it. The problem on the network foundation, like the TCP/IP protocols, where the solution is to incorporate protection into the data flow itself with the sensing abilities of IoT device was addressed. Finally, protecting the link sources would ensure that the device was used safely. In addition, one of the results which [6] referred to was the IoT modules breakdown, which was listed as five main factors as follows:

Equipment or Device: which was incidental to the real device whether it is a sensor, or endpoint or even a washing machine.

Hub or Gateway: A method that can be utilized as a Bluetooth or Ethernet or wireless and etc.

Transport channels or Network, like the satellite and IP networks.

Facilitation: the capacity for transferring data via the gateways and many others, such as processing and analysis.

Application or Consumerization: The willingness of end-users for using details on their mobile phone, and etc.

Subsequently, [6] also lists some threats that were frequently related with IoT, like:

- Intensified Surface attack.
- Systems of legacy.
- Devices that are undetected, prohibited, and invalidated.
- Unauthorized remote access.
- Extensive exposure to sensitive data.

In addition, [6] additionally claimed that these threats must be addressed separately to be related to each other. That's why the CIA can be implemented in every unit. Also the Ubiquitous defense-in-depth strategy was stated in [6] which summarized in the following figure: which will allow the organization to incorporate and track the device security along with the activities and process for the device. It will also include each layer protection and few more technique that would support to accomplish the objectives and identify the threats. Another literature [7], discussed same threats, but also included others that were not addressed in [6]. This was due to the various techniques of identifying threats. So providing security or not it's not the case with [7] instead trying to secure the system in the network base that was part of the things which [6] has discussed, but [7] clarified why it was very important than trying to protect the device with a work strategy or just device on its own [7] stated that securing the IoT device makes it very helpful and appropriate for having an IoT system, like:

- Network Attacks
- Physical Attacks
- Encryption Attacks
- Software Attacks

Hence, to protect the IoT, these risks should be considered and each of the internal risks related to it, as the case of physical attacks on the basis of [7]:

- Node tempering: Which will physically or partially alters the node sensor so that it can provide consistent and direct system access.
- Interference of RF with RFIDs: Through transmitting radio signals to the IoT RFID devices as the DoS attack for generating noise to the device itself.
- Physical damage: damage to the real IoT system and this form of threat was primarily relevant to the protection of the location or building in which the IoT gadget was located [7]. In the other form of threat that is stated by [7] which was a network attack that was similar compared to [6], it actually varies in the form that the attack could occur and how an attack pattern is rendered. So there are more common threats to be prevented in the implementation of the IoT while dealing with particular network attacks.
- Traffic Analysis Attack: It was sniffling in the selected network and then implementing any analytical method utilizing specialized equipments for that purpose that was aimed at one of the things which claimed in [6], that was confidentiality and considered to be a backbone of security.

- Man in the middle attack: the kind of thing described in [6]. Through securing the network and having security protocols to secure the transmutation. However, [7] explicitly specifies what must be happen in the specified attack, how it could occur in this type of attack, where the man in middle was fixed trying to detect the transmission that was going on from the sender and the receiver side of the IoT device.

3.3. Systems review for risk management

Since the risks and vulnerabilities related to the IoT were identified, the literature on risk management can be reviewed. A special publication on the risk management of information technology systems has been published by the National Institute of Standards and Technology or NIST [1]. This publication describes risk management as a three-stage model. It begins by defining the risk, then evaluation, and in the end, by decreasing risk to the appropriate level. It is further noted that risk management helps IT professionals to accomplish the objectives of the organizations by determining a balance within operating costs, control costs, and incident costs.

Therefore, a well-designed risk management method can help to make decisions on the implementation of effective controls. What does the management execute when the residual was better than risk appetite? Hence, the publication noted that management must replicate the risk management iteration till the residual was below or similar to risk appetite. From this, the risk management was concluded as an ongoing procedure that was changing every time. While this work [1] offered the strong basis for process of risk managements, it was very broad and little out of date for present world. NIST, therefore, released a more detailed publication specifically customized to risk management of information security [2]. This work noted that process of risk management consists of four major elements:

- Assess risk.
- Frame risk.
- Monitor risk ongoing.
- Respond to risk.

The four factors and their association with each other are represented in the fig.2.

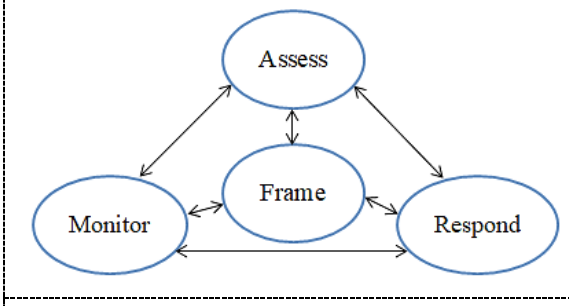

Figure.2. Risk Management Framework

The initial factor of risk management was risk framing. It discusses how organization builds the risk environments. This indicates that an organization must explicitly define a context in where decisions on risk-based are taken. This was not a simple job and involves recognition as following:

- Risk constrains.
- Risk assumptions
- Tolerance to risk.
- Trade-offs and Priorities.

The next factor was assessing risk. It discusses how organizations assess risk between the risk system boundaries. This factor was certainly the significant, and its purpose was to define the followings:

- Risks to the organizations.
- Internal as well as external susceptibilities.
- The risk effect that exploit the vulnerabilities.
- Probability of the attack.

Though, to achieve these objectives, the organizations must define the followings:

- Risk management methods, strategies, and methodology.
- Risk based assumptions.
- Responsibilities and Roles.
- How information on the risk assessment is stored, analyzed, and shared.
- How the risk assessment was carried out.
- Frequency of risk assessments.
- How to obtain information on the threat.

Risk response is the third factor of risk management. How organizations respond to risk after the risk has been defined through a risk assessment was addressed. The purpose of this aspect was to offer a coherent and holistic risk response in accordance with frame of risk. However, this objective could not be accomplished without the followings:

- Develop the alternate risk response protocol.
- Evaluation of the alternative process.
- Determination of the acceptable risk protocol among the sense of risk tolerances.
- Implement the risk responses on the basis of procedures determined.

The final factor of the risk management was continuous monitoring of risk, which discusses how organizations track threats eventually. The purpose of this part was to:

- Assure that the risk responses prepared was well implemented.
- Determining the continued efficacy of risk responses initiatives.
- Detect changes that affect risk.

This work was clearly comprehensive and allows for integrated method to risk management. The International Organization for Standardization or ISO has also published a standard on risk management in the field of data security [10]. When this principle was agreed upon among the network, it lacked a functional feature and did not offer for any deployment, as discussed in [8 & 9]. Nor does it describe an overview of controls presently in effect, as discussed in [11].

TABLE.1. COMPARISON OF RISK MANAGEMENT FRAMEWORKS

Attribute/Framework	NIST SP 800-30	ISO 27005
Method	Tactical	Higher level
Human resource	Not addressed	Explicitly addressed
Information gathering	Interview, Questioner, and document.	Questioner, Interview, and document.
Access	Free access	Paid access

III. ANALYSIS AND DISCUSSION

In the literature review, numerous flaws and problems associated with IoT systems were discussed and various risk assessment systems and their elements were addressed. The discrepancies and similarities among these systems were analyzed. On the basis of the analysis, that NIST 800-30 was essentially a management framework that blends into a technology-related context and was therefore more technical. Hence, ISO 27005 was better suited to management of higher-level activities, as it encompasses systems, people and technologies. However, these two systems were applicable to several organization that means that they were very generic for anything like an IoT unit.

IV. PROPOSED SOLUTION

After evaluating the various existing risk management systems and analyzing the findings, the following are proposed, rather implementing the common risk management system to every current IoT devices.

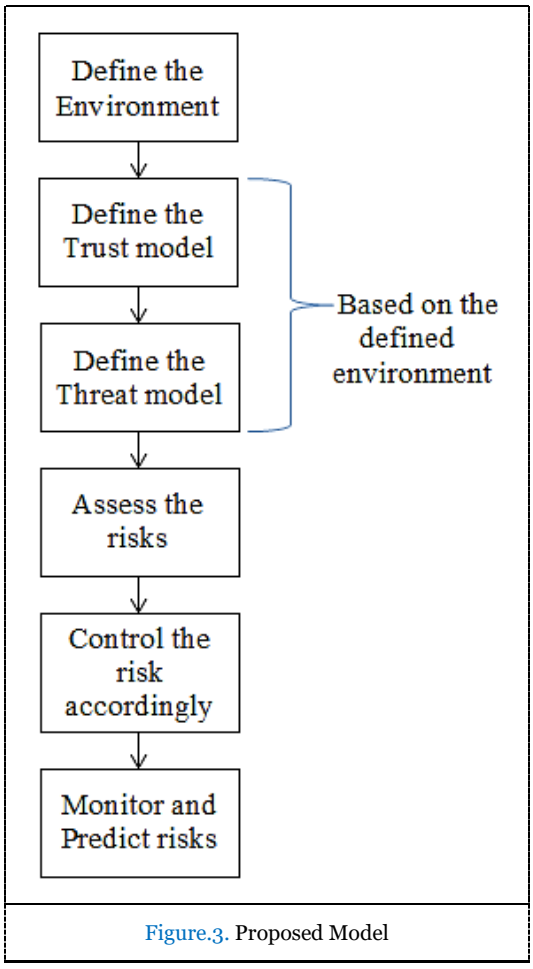

Figure.3. Proposed Model

The process of risk management should be incorporated into the life-cycle production

of IoT devices itself. The IoT gadget manufacturer must incorporate an acceptable risk assessment process depending on the design of IoT devices. In this manner, the execution of risk managements can be judged on basis of the device's performance. i.e., the application of risk managements for the surveillance cameras can be done else as relative to the smart refrigerator. That was precisely why the existing risk management systems could not be applied on every IoT applications, because the present risk management framework seem to take the overview of and thus over-generalize risk management. It is comprehended that the IoT system is highly varied. It is also not necessary to implement the similar risk managements system for a microcontroller and the smart lock that was liable for protecting the entire buildings. Generally, trust and threat model vary considerably within IoT device, and thus, the standard risk managements system for each of them cannot be implemented. The risk managements process must be managed during the construction stage of the IoT facility for achieving the most successful and reliable process of risk management feasible. In the figure.3, the model that functions as a module must be incorporated into the life cycle of the IoT system.The initial step in the developed framework was to define the framework or environment in which the IoT interface is supposed to function. Further, an anticipated outcome of this process was a good description of the functionality and shortcomings of the system. In addition, the extrinsic and intrinsic values of the device must also be specified. The purpose of this step was to setup the way for the further steps.

The next step was the concept of the trust model. The specified model must comprise the hardware, software, and information on which a security of the device depends in relation to the defined context. The purpose of this step was to identify which components can be trusted by the IoT system and, thus, to identify what the components of untrusted are, which was important for the following step.

The next step was the concept of a threat model. The identified threat model was intended to recognize vulnerabilities related to the IoT system. It should also classify main threats which might target these vulnerabilities. In addition, the threat model should describe the possible collection of steps that an attacker could take to breach the device. Eventually, the threat model must recognize the effect as soon as the identified threat exploits the vulnerability. All of this must be performed in the sense of the given context. The purpose of this process was to collect the details required for the risk assessments to be carried out in the next step.

The next step, and also the significant one, was to evaluate the risk in line with the performance of the last steps. Risks assessment was a two-step method of risk recognition and risk assessments. The outcome from the past phase to define the possible risk associated in IoT device will be used. Then, after the risk has been established, it will reflect it in a qualitative or quantitative manner depending on the likelihood of occurrence and its effect after it has occurred. The purpose of this process was to direct the decision-making process for reacting to these threats, which would occur in the following step.

The following step was entirely about the use of controls to minimize risks to the optimal level. Though, it was important than a feasibility analysis in advance, particularly in the context of IoT was performed, since the lot of IoT gadgets were inexpensive. The implementation costs and the cost of the effect should be covered by the safeguards applied. The output of this step was an IoT system that was already handled in relation to vulnerabilities and risks and was able to be implemented without further risk control from user. This could be the last process in few situations where it was not possible to install patches and manage IoT systems.

The last step in model was to monitor existing threats and predict potential risks. Residual risks left out of the previous phase should be tracked to assure that the risk appetite was still smaller. Still need to predict potential threats, since technology was constantly changing and new vulnerabilities were developing day-by-day. IoT systems must also be maintained update in order to ensure an acceptable security level. This may be achieved by a range of means, such as patches for security and upgrades. Though, it was important to comprehend that this process will not be possible for few IoT gadgets, hence the expense of this process might be much higher than benefit of an IoT product. In the light of

the knowledge given in the earlier steps, the feasibility review may determine either or not it could be necessary to implement this process. Unlike earlier steps, this process was an on-going one.

The proposed approach has many benefits over the existing risk management systems, few of them are:

- Appropriate for all IoT devices: Since the proposed approach is incorporated into the product development of all IoT devices, every system will have its own risk management system if required.
- Minimum charges from the User: The model required that risk management would be completely managed by the manufacturer. Therefore, no extra risk assessment can be added to the system by the end user.
- Very effective: Because the device manufacturer knows the best about the device, it could be ensured that the risk management frameworks offered by the manufacturer was very efficient.

V. CONCLUSION AND FUTURE WORK

IoT devices have been a critical part of our lives. However, due to their design, certain IoT devices are vulnerable to different forms of threats, which make it much more difficult to protect them. The importance of handling these vulnerabilities correctly and its built-in threats is thus greatly emphasized. Hence, a new approach to risk management, which is more suitable, because it is uniquely customized to each IoT system, needs minimum overhead from the user and is typically more efficient. In future, the problems of added cost of development will be addressed and also an effort to reduce it.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

Full Degree Study, Stipendium Hungaricum Scholarship, Reference Number: SHE-19575-004/2020, Dated: 29.07.2020.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The author would like to thank Faculty of Informatics, Eötvös Loránd University for providing support while carrying out this work.

REFERENCES

[1] A. Feringa, A. Y. Goguen, and G. Stoneburner, (2002). Risk management guide for information technology systems. Special Publication-800-30.

[2] R. S. Ross. 2011. Managing Information Security Risk: Organization, Mission, and Information System View (No. Special Publication (NIST SP-800-39).I. S. Jacobs and C. P. Bean. Fine particles, thin films and exchange anisotropy. in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic.

[3] Evans, D. (2011). The internet of things: How the next evolution of the internet is changing everything. CISCO white paper.

[4] R. Shapaval, and R. Matulevičius. (2018). Towards the Reference Model for Security Risk Management in Internet of Things. International Baltic Conference on Databases and Information Systems.

[5] J. Tong, W. Sun and L. Wang. (2013). An information flow security model for home area network of smart grid. IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems.

[6] A. Jha, and M. C. Sunil. (2014). Security considerations for Internet of Things. L&T Technology Services.

[7] I. Andrea, C. Chrysostomou and G. Hadjichristofi. (2015). Internet of Things: Security vulnerabilities and challenges. IEEE Symposium on Computers and Communication (ISCC), Larnaca.

[8] A. Asosheh, B. Dehmoubed, and A. Khani. (2009). A new quantitative approach for information security risk assessment. 2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT).

[9] A. Ekelhart, S. Fenz, and T. Neubauer. (2009). Aurum: A framework for information security risk management. 42nd Hawaii International Conference on System Sciences, HICSS '09.

[10] ISO/IEC. ISO 27005 information technology security techniques information security risk management, 2008.

[11] N. Al-Safwani, S. Hassan, and N. Katuk. (2014). A Multiple Attribute Decision Making for Improving

Information Security Control Assessment. Int. J. Comput. Appl.

[12] G. Soos, D. Kozma, J. Nandor, Ferenc, and P. Varga. (2018). IoT Device Lifecycle – A Generic Model and a Use Case for Cellular Mobile Networks.

[13] L. F. Rahman, T. Ozcelebi, and J. Lukkien. (2018). Understanding IoT systems: a life cycle approach. Procedia Comput. Sci.

[14] G. Ambika, and P. Srivaramangai. (2018). Review On Security In The Internet Of Things. Int. J. Adv. Res. Comput. Sci., Vol.9, No.1.

[15] M. A. Razzaq, S. H. Gill, M. A. Qureshi, S. Ullah. (2017). Security Issues in the Internet of Things (IoT): A Comprehensive Study. Int. J. Adv. Res. Comput. Sci. Appl., Vol. 8, No. 6.

Cite this article as: Mugesh Ravi. A Survey on Security Risks in Internet of Things (IoT) Environment. J. Comput. Sci. Intell. Technol. 2020; 1(2): 01–08. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 2, pages: 09–14, September 2020

A Survey on Cloud Computing for Information Storing

Sathiyasheelan Ravichandran

Faculty of Business Informatics, Riga Technical University, Riga, Latvia.

***Corresponding Author: rsathiya2196@gmail.com*

Received: *17.06.2020,*
Revised: *10.07.2020,*
Accepted: *10.08.2020,*
Published: *31.09.2020*

DOI:
10.53409/mnaa.jcsit20201202

Abstract: Cloud computing is a technique for storing the information virtually. It may comprise of database, storage, tools, servers, networking and software services. It deals with virtual storing and retrieving of data from anywhere by the help of internet. This paper uses cloud computing technique in education for uploading study materials, videos, sharing information to the students and for conducting tests. The symmetric key encryption technique is used in this concept, where one key is utilized for both decryption and encryption. The advanced encryption standard algorithm (AES) was used for securing the data in the cloud, where it gives high security and faster execution time. This technique is mainly based on improving the concept of virtual classroom by using cloud computing.

Keywords: *AES algorithm, cloud computing, E-learning, storage, virtual classroom*

I. INTRODUCTION

Cloud computing is a biggest platform for all services providers over the internet. The main resources of the cloud provide storage, server, database, networking and software for all the tools and applications. Cloud is a major platform for the educational field for online learning, smart technology and networking, for each process of services and local storage devices, hard drive it will report the database for the verification. By using cloud services for a college management, business process, marketing fields. The reason for cloud computing option for people and business including cost saving, increased the product productivity, speed and efficiency automatically increases, performance of the services in all platform, security is highly effective by using different algorithm based on the applications. In cloud computing the information is accessed virtual space with the help of remote. Business aspects of all companies provide cloud service user can enable to save applications and files on remote servers and then access every data through internet.

In recent years, cloud computing has demonstrated a variety of points of concern over traditional figuring models to be mainstream. Average desirable conditions with mobility, the convergence of adaptability, energy conservation, and cost savings. Cloud computing has been introduced as a general term for describing a set of advanced on-demand services for computing offered by commercial providers like Google, Amazon, and Microsoft. It means a computing infrastructure model is considered as a "cloud," from which organizations, industries, and people accesses applications on demand across the world. The fundamental concept beneath this technology was providing storage, computing, and software "as a service". National Institute of Standards and Technology (NIST) describes Cloud computing as a "model for enabling universal, convenient, and on demand access of network to configurable computing resources in a shared pool which could be quickly provided and delivered with minimum managerial efforts or service provider interactions".

Cloud is an on-demand advance model for the IT world, constantly taking into account simulation and cloud computing developments. Cloud computing involves both the applications represented as services through Internet and hardware and software of the device used in the data center to offer the service. The systems

themselves were known as IaaS (Infrastructure as a Service), Software as a Service (SaaS), and PaaS (Platform as a Service) for describing their service.

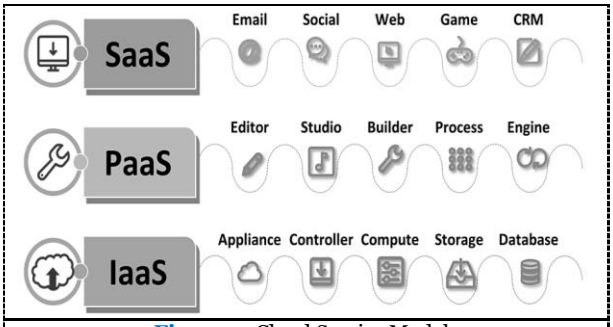

Figure.1. Cloud Service Models

Cloud computing process data from the devices from around the all storage of data based on the virtual platform services providers. The internet cloud services data and application and work are accessible from any devices with could link to the internet. Cloud computing have different types of user can approaches the cloud services like public cloud, private cloud and hybrid cloud. Public cloud service provides their service to the user through the internet for free. Private cloud service provides their service to the business aspects of companies for certain number of people. The combination of public cloud and private cloud of both elements are known as hybrid cloud, these services provide system of network that supply hosted services.

Although there are no official recommendations for the use of cloud-based systems in the education field, an inquiry to define and establish a reliable platform as a structured method in higher education is desirable as well as an area of interest and study. There is no question that teaching and learning strategies focused on personal computer equipment may pose a serious risk of exacerbating the divide among students who have access to different resources and technologies and students who cannot endure the same.

II. RELATED WORK

2.1. Cloud Computing – An Advanced E-Learning Platform of School Education

Cloud computing trends are use to develop the e-learning services. A platform of e-learning is fully concentrated on student education system and their knowledge of the upcoming technologies. The e-learning system estimate the need of new platform that are depend upon the techniques like mobility, flexibility, individualization availability, and openness of education. Cloud technology are used in facilities and computer resources are available on the web service user like platform as a services, software as a services, hardware as a services, infrastructure as a services, communication as a services. The user accesses the online cloud computing by using certain equipments like laptop, net book, and smart phone. The advantages of these techniques reduce the hardware requirements and cost. Implementation of cloud computing technology is developing the modern education system it can use the main concepts of e-learning techniques to learn the current technology for improving the education resources.

2.2 E-Learning in a Cloud Computing Environment

Cloud is a growing technology in every sector the storage services is include based on the different types of services and the main goal of the cloud to develop the many education institutions with high ability for infrastructures and resources to execute the e-learning system. The concept of the paper explains the advantages and limitation of the cloud e-learning. Cloud computing is the growth future e-learning across institutions of worldwide services. Educational institution is focusing on offering students, faculties and administrations with capability to improve the modern mobile world.

2.3. The Application of Cloud Computing In Education Informatization

The application research of cloud computing in educational information, the tradition of technologies based on computers include the network storage, virtualization, distributed computing play a major role to share the current status of the cloud service, parallel commuting techniques is used to delivered the information on the same time to the users, network technology is a main process for sharing a data with fast and secure manner and automation techniques is initially the process is start until the last the same process. The proposed method of cloud computing is used by IBM and GOOGLE platform. The cloud computing was significance for developing information technology in field of education. The cloud is a convenient and cheap for more data processing

of modern education system. The level of education technology will be enhanced the information sources in future.

2.4. A Proposed Model for Education System using Cloud Computing

Education and cloud computing combine to give a quality type of education under the information technology develop the economic growth. The faculties are ready to develop an informatics future student under the backbone of cloud services providers like sales force. The information sharing techniques make a student's more knowledgeable atmosphere. Many services providers offer various cloud based application for user who can work simply for academia of cloud computing. The e-learning system facing teaching challenges of cloud computing. The propose model of cloud computing the teachers and students could share the course programs through the cloud and students can update their examination, materials and assignments. The attendance records are also stored in cloud services. Cloud computing education was to minimize the difficulties in teaching infrastructure.

2.5. E- Learning Based on Cloud Computing Technology

Cloud computing is a growing rapidly, with application including the area based on education. The telecommunication devices like desktop, laptops, tablets, mobile, music players it make a product to distributed compare to any centralized entity. E-learning system is usually require many hardware and software resources offer the cloud computing facilities to achieve flexibility, efficiency for good process e- learning services. Cloud e-learning is an approach for selecting the cloud as it promise very clear advantages. Cloud support digital services for storage data to distributed the access centralized system uses of application and resources for several cloud applications. By using the cloud platform is less amount of cost and the data sharing under safe and secure process of cloud services providers increase the learning facilities.

2.6. AES Algorithm Based Approaches

This segment highlights the significance of advanced encryption standard (AES) technology, as it is the key algorithm method used in the conducted study of both the encryption process and the decryption process. A variety of papers was evaluated on the basis of a testing methodology in which the AES algorithm is used in a number of cloud services related applications in which other algorithms are effectively compared.

Abha Sachdev and Mohit Bhansali (2013), with the monumental growth of sensitive cloud data, cloud protection has become more critical than any time in the modern past. Cloud information and software live in highly adaptable storage systems and can be accessed all over the world. Unfortunately, the advancement of cloud clients was combined with the emergence of malicious behavior in the cloud. An ever-increasing number of bugs have been discovered and new security alerts have been regularly issued. A vast number of clients access the Cloud for diverse reasons and thus need incredibly safe and resourceful services. The development of the cloud, especially in the increasing scope of apps, includes far more protection and validation.

Roshani Raghatate et al., (2014), Adding AES to information security provide benefits of less memory usage and less time to measure as compared to alternative algorithms. About the fact that cloud infrastructure has its own security features, the client should select architecture depending on its security. Each cloud service has its own set of rules, pricing, capability, support as well as other essential features. The primary consideration dealt with in this proposal was the encryption concept for protecting data by making it inaccessible to everyone.

A.M Abdullah, (2017), AES algorithm was one of the familiar and commonly utilized symmetric block cipher algorithms used worldwide. This algorithm got the individual basic framework for encrypting and decrypting confidential data and was utilized in hardware and software worldwide. It was very challenge for hackers to acquire the original data when encrypted with an AES algorithm. No proof was there to date that this process of algorithm could be solved. AES could accommodate three different sizes of keys, such as AES 128, 192, and 256 bits, and both of these 128-bit block sizes.

III. CLOUD SERVICE MODELS

Cloud services are provided as IaaS, PaaS, and SaaS. It is a combination of many techniques, comprising distributed and grid computing, and

service delivery network by using Internet. The public cloud platform is very complicated when contrasted with a conventional data center environment. Based on the model of Cloud computing, an institution or organization gives up direct access to significant features of security, provides a higher level of trust over the Cloud provider. The three different types of service models in cloud computing are:

3.1 IAAS

Infrastructure as a service (IAAS) is a virtual resources that includes virtual servers, networks, processing power, hardware and storage. These services can be accessed by cloud computing platform based on pay as per usage of services. This service reduces the cost and complexity of buying the software and managing the servers rather than using as we need. The services are highly scalable and flexible and also maintain the system backup and developing measures.

3.2 PAAS

Platform as a service (PAAS) was a cloud computing service model which provides storage, database management system, hardware, networking, servers, middleware, software, operating system, application framework, programming languages and other development tools to users for creating applications. These services can be easily accessible by cloud platform services based on pay for usage. This is highly scalable and it makes easy for the customers to create, run, manage, test and deploying the web applications.

3.3 SAAS

Software as a service (SAAS) was an on-demand software that provides software to the users by accessing through internet. This service is cost effective and the users no need to pay money for purchasing and installing the software because it works on monthly subscription of software. The software is provided by third party vendors as a host application to the customers based subscription of products. This model can be used in business applications, customer relationship management, document management, social networking and mail services.

3.4. Selection of cloud deployment model

A cloud computing based solutions can be used in following ways,

Community Cloud: This uses a framework with network for its assurance just like a grid.

Public cloud: This kind of solution is given by the Google, amazon where the data must be ensured in a private framework with an assurance that it is ensured there also it relies upon pay start and demands more prominent security from a broad amounts of vindictive groups.

Private Cloud: This kind of feature is required in private affiliations and government firms, and besides where the data stored needs more care and should be managed carefully. These can develop their own particular plan of standards in a cloud where simply the administrators of that affiliation can get to.

Hybrid Cloud: It is mix of the two clouds described above i.e. Private cloud and Public cloud.

Cloud computing has been used in a variety of applications globally. The research work was in the domain of cloud computing, for which technology was necessary to be used for technological development in the education application. Several companies or organizations now have minimum one program or half of their computing technology over the cloud. Cloud environments are ripening and, in few cases, getting more nuanced. Although 43 percent use just hybrid cloud and 12 percent just use multi cloud, 30 percent use both.

IV. SYSTEM DESIGN

The architecture diagram (see fig 1) explains the concept of cloud computing based virtual classroom technique. This system helps to improve the communication between the students and the staff by storing and retrieving all the information from the cloud. The cloud computing contains all the details of staffs, students, study materials, videos and institution related information. Here, the cloud is connected with staff and students, where staff is responsible for uploading study materials, videos, conducting tests, viewing student performance and sharing general information among the students. Students are responsible for taking tests from anywhere, viewing videos, materials and information, communicating with

friends and can do the project by the use of cloud computing.

Cloud computing based virtual classroom technique makes the students to gain knowledge from anywhere and from any device. It helps the staff to manage more number of students in a single time and can view the student's performance from anywhere and from any device. This technique does not need any storage device for storing the details, where it has the capacity of storing the information virtually in cloud. This paper is based on symmetric key encryption, it contains single key for both encryption and decryption techniques. The algorithm used for this technique is advanced encryption standard algorithm (AES). AES is commonly typically used for cloud protection. AES was proposed in order to replace DES in functional implementations. AES was adopted by the NIST on 26 November 2001. AES was a symmetric-key algorithm that means that the single key was used for both data decoding and data encoding. AES was also called RIJNDAEL, named after the names of its inventors John Daemen and Vincent Rijmen. AES was unpredictable and depended on the duration of the key. AES used 10 series for 128-bit keys, 14 series for 256-bit keys and 12 series for 192-bit keys. Both of the series uses an alternative 128-bit series key, which was predicted from the original AES key. AES algorithm provides high security, allows less space and quicker execution time than other algorithms. This algorithm encrypts and decrypts data using cryptographic bit keys. This guarantees high protection for data in the cloud world.

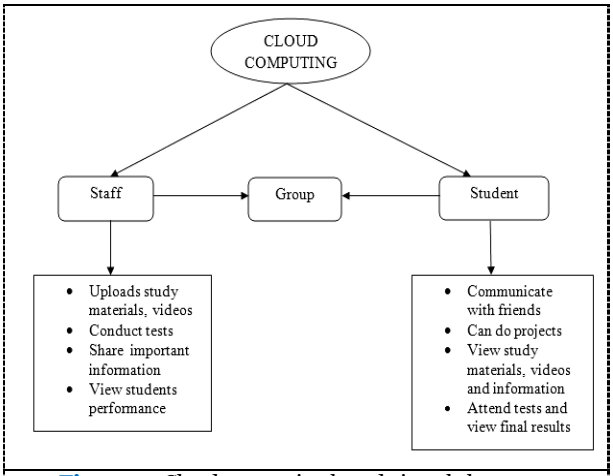

Figure.2. Cloud computing based virtual classroom technique

AES works well in both the software and hardware processes under a wide range of circumstances. This includes 8-bit and 64-bit systems and DSPs. Its underlying parallelism enables the optimal use of processor resources contributing to execution of large-scale programming. This algorithm got fast key configuration time and improved key operation. It needs less memory for use, making it suitable for small space scenarios. The structure had a tremendous ability to benefit from parallelism at the stage of teaching. There were no critical bad keys to AES. It supports both block size and key size which were 32 multiples (more notable than 128-bits). A numerical measurement of the ciphertexts was not possible despite a huge number of experimental had been carried out. No differential and linear cryptanalysis attacks on AES have yet been developed.

Proper convergence of face-to-face and interactive learning is needed to ensure the continuity of higher education through the use of digital skills. A proper balance of digital life and physical life will yield enormous benefits for the economy and society. Future learning education would focus more on artificial learning, virtual machines/online computers, or cloud-based information storage. Real-life and virtual learning will complement one another, and consumers would have remote access to apps independent of time and physical spaces.

Transition to virtual learning was an interim solution to the coronavirus pandemic. What is needed is rethinking, creativity, automation, and quality management, with an emphasis on learning digitalization. The design of the layout of the courses and the mode of delivery in higher education must be based on scenario planning. Instructional design is a common concept coined in virtual learning that ensures online learning is most successful. This examines the ability to apply modern technologies to the immersive and innovative implementation of e-classes. To sum up, successful digital learning involves imagination, ingenuity, and the design of innovative technologies to involve students before and during lectures on the digital platform. A constructive strategy is required to share the expertise and experience of workers with varied needs.

V. CONCLUSION

Cloud computing based virtual classroom technique is mainly based on improving the communication between the staff and the student by storing and retrieving all the information's from cloud. Cloud computing helps the people to store large amount of data virtually and allows the people to retrieve the data from anywhere and by any device. Advanced encryption standard (AES) algorithm is used in this concept, where it provides high security, faster execution time, and high encryption capacity and requires low amount of memory usage. This system helps all the educational institution for storing, retrieving and viewing all the students and staffs performance virtually by using cloud computing technology.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The author would like to thank Faculty of Business Informatics, Riga Technical University for providing support while carrying out this work.

REFERENCES

[1] Ishaq A. and Brohi M. N. Cloud Computing in Education Sector with Security and Privacy Issue: A Proposed Framework. International Journal of Advances in Engineering & Technology vol. 8, no. 6, pp. 889-898.

[2] Engr. Ali Ahmed, Huma Ali Ahme. A Proposed Model for Education System Using Cloud Computing.

[3] Hongyu Pei Breivold and Ivica Crnkovic. (2014). Cloud Computing Education Strategies. 978-1-4799-4970-0.

[4] Majid Shirzad, Ali Hoseinpanah, Mehdi Ahmadipour, Hojat Rahimi. (2012). E-Learning Based on Cloud Computing, of20121ntemational of Cloud Computing, Technologies, Applications & Management 978-1-4673-4416-6.

[5] Bo Wang, HongYu Xing. (2015). The Application of Cloud Computing in Education Informatization, 978-1-4244-9763-8.

[6] Shaoyong Chen, Min Lin, and Huanming Zhang. (2019). Research of mobile learning system based on cloud computing. International Conference on e-Education, Entertainment and e-Management, 978-1-4577-1382-8.

[7] Utpal Jyoti Bora, Majidul Ahmed. (2013). E learning using Cloud Computing International Journal of Science and Modern Engineering (IJISME) ISSN: 2319-6386, Volume-1, Issue-2.

[8] Shyshkina Mariya. (2016). Cloud computing – an advanced e-learning platform of school education, 14th International Conference on Interactive Collaborative Learning, 978 -1-4577-1747-5.

[9] Mohammed Ketel. (2014). E-Ieaming in a Cloud Computing Environment, 978-1-4799-6585-4.

[10] Mohssen M. Alabbadi. (2011). Cloud Computing for Education and Learning: Education and Learning as a Service (ELaaS), 14th International Conference on Interactive Collaborative Learning, 978 -1-4577-1745.

[11] Roshani Raghatate, Sneha Humne, and Roshna Wadhwe. (2014) A Survey on Secure Cloud Computing using AES Algorithm, IJCSMC, Vol.3, Issue.12, pg.295 – 301.

[12] Abha Sachdev and Mohit Bhansali. (2013) Enhancing Cloud Computing Security using AES Algorithm, International Journal of Computer Applications (0975 – 8887) Volume 67– No.9.

[13] Ako Muhamad Abdullah. (2017). Advanced Encryption Standard (AES) Algorithm to Encrypt and Decrypt Data, Cryptography and Network Security.

Cite this article as: Sathiyasheelan R. A Survey on Cloud Computing for Information Storing. J. Comput. Sci. Intell. Technol. 2020; 1(2): 09–14. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 2, pages: 15–25, September 2020

A Hybrid Genetic-Neuro Algorithm for Cloud Intrusion Detection System

Suresh Adithya Nallamuthu

Gdansk University of Technology, Gdansk, Poland.

***Corresponding Author:** sureshadithya1991@gmail.com*

Received: *10.06.2020,*
Revised: *10.07.2020,*
Accepted: *19.08.2020,*
Published: *31.09.2020*

DOI:
10.53409/mnaa.jcsit20201203

Abstract: The security for cloud network systems is essential and significant to secure the data source from intruders and attacks. Implementing an intrusion detection system (IDS) for securing from those intruders and attacks is the best option. Many IDS models are presently based on different techniques and algorithms like machine learning and deep learning. In this research, IDS for the cloud computing environment is proposed. Here in this model, the genetic algorithm (GA) and back propagation neural network (BPNN) is used for attack detection and classification. The Canadian Institute for Cyber-security CIC-IDS 2017 dataset is used for the evaluation of performance analysis. Initially, from the dataset, the data are preprocessed, and by using the genetic algorithm, the attack was detected. The detected attacks are classified using the BPNN classifier for identifying the types of attacks. The performance analysis was executed, and the results are obtained and compared with the existing machine learning-based classifiers like FC-ANN, NB-RF, KDBN, and FCM-SVM techniques. The proposed GA-BPNN model outperforms all these classifying techniques in every performance metric, like accuracy, precision, recall, and detection rate. Overall, from the performance analysis, the best classification accuracy is achieved for Web attack detection with 97.90%, and the best detection rate is achieved for Brute force attack detection with 97.89%.

Keywords: *Intrusion Detection, Cloud Computing, Genetic algorithm, Back Propagation Neural Network, CIC-IDS*.

I. INTRODUCTION

Cloud computing has been introduced as a general term for describing a set of advanced on-demand services for computing given by commercial suppliers like Google, Microsoft and Amazon. It means the computing infrastructure model was considered as the "cloud," from which organizations, industries, and people accesses application on demand worldwide. The fundamental concept beneath this technology was providing storage, computing, and softwares "as a service." [1]. National Institute of Standards and Technology (NIST) describes Cloud computing as the "model for enabling universal, convenient, and on demand access of network to a configurable computing resources in a shared pool which could be quickly provided and delivered with minimum managerial efforts or service provider interactions" [2].

Cloud services were provided as Infrastructure-as-a-service (IaaS), Platform-as-a-service (PaaS), and Software-as-service (SaaS). It is a combination of many techniques, comprising distributed and grid computing, and service delivery network by using Internet. The public cloud platform is very complicated when contrasted with a conventional data center environment. Based on the model of Cloud computing, an institution or organization gives up direct access to significant features of security, provides a high level of trust over the Cloud provider [3-5].

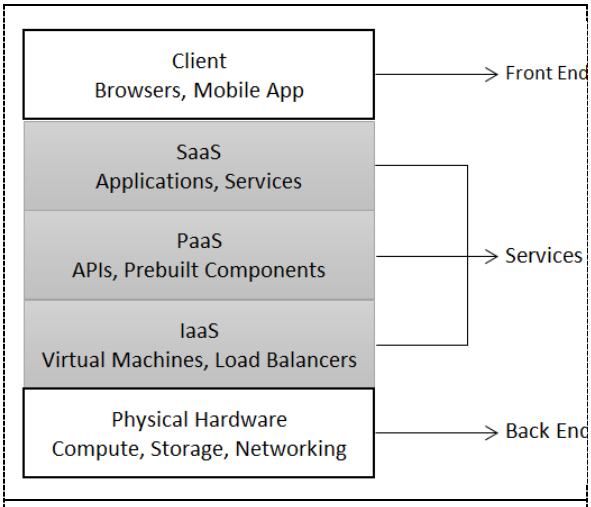

Figure.1. Cloud Computing Stack

In the Cloud system, distributed and shared resources makes it challenging to create a security system to ensure data privacy and security. Because of open issues, no Cloud service provider (CSP) accepts their users to use IDS or security techniques expanding into the management service layer behind virtualized Cloud instances. Users might not know about specific security-issues, vulnerabilities, or malware details. For example, the intruders might have the option to obtain the information of Cloud accounts and set up a kernel-level rootkit through back-channel [6]. Intrusions on "physical level" are like extracting the RAM of the virtualized host or undermining the virtualization layers were aware to the network. Indeed, the host system offering the information cannot be reliable entirely once the CSP holds the physical resource. CSPs regularly setup a SLA (Service Level Agreement) to feature the privacy and security of the relevant services [7].

Though numerous security complexities arise with the variation to this computing model, including IDS, however, the significant development of the data security innovations recently, attacks and intrusions keep on defeating existing IDSs in the Cloud environment. A disastrous DDoS attack recently has taken down over 70 significant Internet services, including Github, Amazon, Paypal, Twitter, and so on. Attackers have exploited IoT and Cloud Computing technologies to produce a large volume of traffic attack over 665 GB/s [8-9].

Intrusions and attacks have become a challenge to the existing Cloud IDSs through enormous amount of network traffic data, complex and dynamic actions, and new attack classes. IDS for Cloud must analyze network traffic data with vast volumes, identify the new attack actions effectively, and obtain higher accuracy with a low error rate [10].

In this research, a neural network-based technique Back Propagation Neural Network (BPNN) is proposed to classify the detected intrusions in the cloud environment and for detecting the intrusions the Genetic Algorithm (GA) is proposed. From the proposed dataset, the data are preprocessed and forwarded to detect the present anomalies by using the genetic algorithm. This detection enables the model to recognize and blocks suspicious data while granting accesses to general data. The anomaly data are then classified by using the BPNN classifier to detect each attack type present in the dataset.

II. RELATED WORK

IDS is a significant security tool used to secure the resources of the cloud. However, IDS frequently experience poor detection accuracy because of composed attacks like DDoS. Even though specific works have limitations and lack of methodology to decide a proper time to exchange attack data between nodes in the distributed IDS. In this way, N M Ibrahim and A Zainal proposed the distributed IDS that used an algorithm called binary segmentation change point detection, toward addressing the proper time frame to forward attack data to distributed IDS nodes and utilizing parallel Stochastic Gradient Descent with SVM (SGD-SVM) for accomplishing the distributed identification. This model was experimented in Apache Spark utilizing the NSL-KDD dataset [11].

The increasing demand for cloud computing causing it inclined to different attackers impacting the integrity and privacy of the information stored in cloud. The efficient IDS composed of a minimum time-compelling algorithm with minimum space complication and high accuracy. To make this, the total features were decreased while preserving less data loss. Partha Ghosh et al. proposed a feature selection model, in which the features were chosen based on mutual data gain between related attributes. To accomplish this, initially, the attributes as per the correlativity were grouped. Hence, from every group, the attributes with the most elevated mutual data gain in their relevant group were chosen. This resulted in a minimized feature set that provided fast learning and hence delivered great IDS that will secure the information in cloud [12].

Identification of attacks and intrusions through illegal users is perhaps the greatest challenge for both cloud service providers and

users. Multi-Layer Perceptron Neural Networks and Particle Swarm Optimization were used by A.S Saljoughi et al. for identifying the attacks and intrusions in cloud computing. For optimizing the neural network, a combination of the neural network with the PSO was used to extricate optimal weights and attempted to decrease the time complications via training the network with random weights. The K-fold technique was utilized and selected a random group for result analysis. The model is assessed with KDDcup99 and NSL-KDD datasets [13]. Another combination of algorithms like MLP network, artificial bee colony, and fuzzy clustering were presented by B Hajimirzaei and N J Navimipour for detecting intrusions in the cloud environment [14].

Mohamed Idhammad et al. presented a distributed IDS based on machine learning for the Cloud environment. The model was developed to be embedded in the Cloud, along with the cloud provider's edge network elements. This enables intercepting incoming network traffic to the physical layer's edge network routers. The captured information was then preprocessed and forwarded to the initial detection of the anomaly process utilizing a Naive Bayes method. This enables detecting and blocking speculated traffic while granting normal traffic access. The speculated traffic at every network router part was synced to a main server. The Random Forest classifier was utilized for classifying the network traffic information accessible on the server and detects each attack type. This IDS model was performed on the Google Cloud and evaluated by the dataset CIDDS-001 [15].

A hybrid model composed of FCM and SVM was proposed by A.N Jaber and S.Ul Rehman. This hybrid technique was used as an IDS model to identify intrusions and attacks in a cloud computing environment. This model was separated into three stages. The primary stage presented the FCM clustering module that was utilized to divide huge dataset into small cluster for allowing the SVM to learn viably in an ideal way. The fuzzy model improved the SVM's performance. In the next stage, various SVM models were trained as per the allocated cluster values. Finally, the fuzzy aggregation model combined the results of the hypervisor inspector [16].

Edge computing expands conventional services of cloud to the network edge, and the exceptionally heterogeneous and dynamic condition at the network edges creates the security of networks circumstance confronting extreme difficulties. H.Yin et al. analyzed the improved k-dependency bayesian network technique that defined the trust relations between system elements and decreased the difficulty of the BN structure by lessening the coordinated weak dependence edges. By presenting a virtual augmentation technique and the maximum a posterior (MAP) criterion for small category samples, this classification model for intrusion detection based on improved KDNB was developed. The results were assessed using just 10% of the KDDCup99 dataset. This model solved the issues of poor stability and low detection accuracy [17].

III. PROPOSED METHODOLOGY

In this proposed model, the Genetic algorithm and BPNN classifier algorithms are used for detecting the intrusions and attacks in the cloud computing environment.

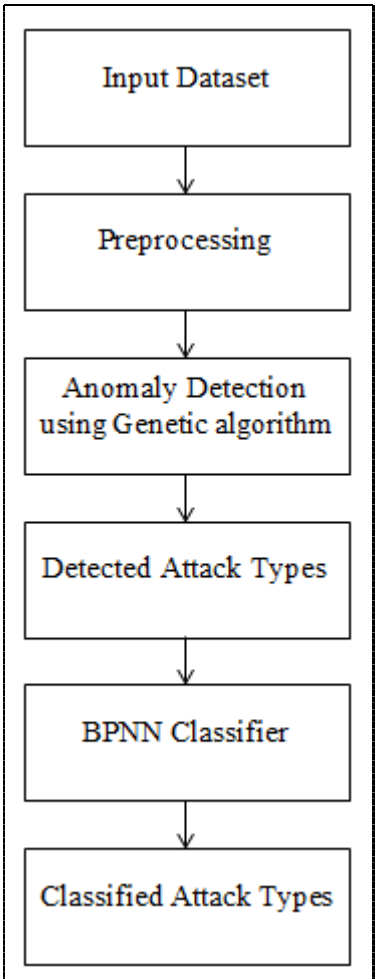

Figure.2. Proposed Method

Initially, the dataset is given as input to the system and preprocessed. For detecting the attacks and intrusions, CICIDS 2017 dataset is used. The preprocessed data forwarded to the initial anomaly detection using a GA, and the

results from the GA are fed to the BPNN classifier. These results from the genetic algorithm are used to detect the intrusions and attacks present in the dataset as malignant and normal data. The BPNN classifier is used to classify the types of attacks.

3.1 Genetic Algorithm

The GA updates the population of individuals iteratively. By using fitness function, the individuals are assessed during every iteration process. A new population generation is acquired from the current generation by probabilistically choosing fitter individuals. Some of the individuals were allowed unchanged to the next generation. Others are referred to genetic operators like mutation and crossover to make new offspring. g is the current generation; n was the total individual in the populations; Z was the population fraction to be substituted by a crossover in every iteration, and μ was the mutation rate.

//Initialize generation
g=0;
P_g = n randomly produced individuals population;
//Calculate P_g
Calculate fitness (f) for every $f \epsilon P_g$;
Do
{ //Create generation g+1
//1. Copy
Choose (1-Z) × n individuals of P_g and insert into P_{g+1};
//2. Crossover
Choose Z × n individuals of P_g; pair them up; generate offspring; insert offspring into P_{g+1};
//3. Mutate
Choose μ × n individuals of P_{g+1}; invert a randomly chosen bit in each;
//Calculate P_{g+1}
Calculate fitness (f) for every $f \epsilon P_g$;
//Increment
g= g+1;
}
while the fitness of the fittest individual in P_g is not enough high;
return the fittest individual from P_g;

3.2 Back Propagation Neural Network

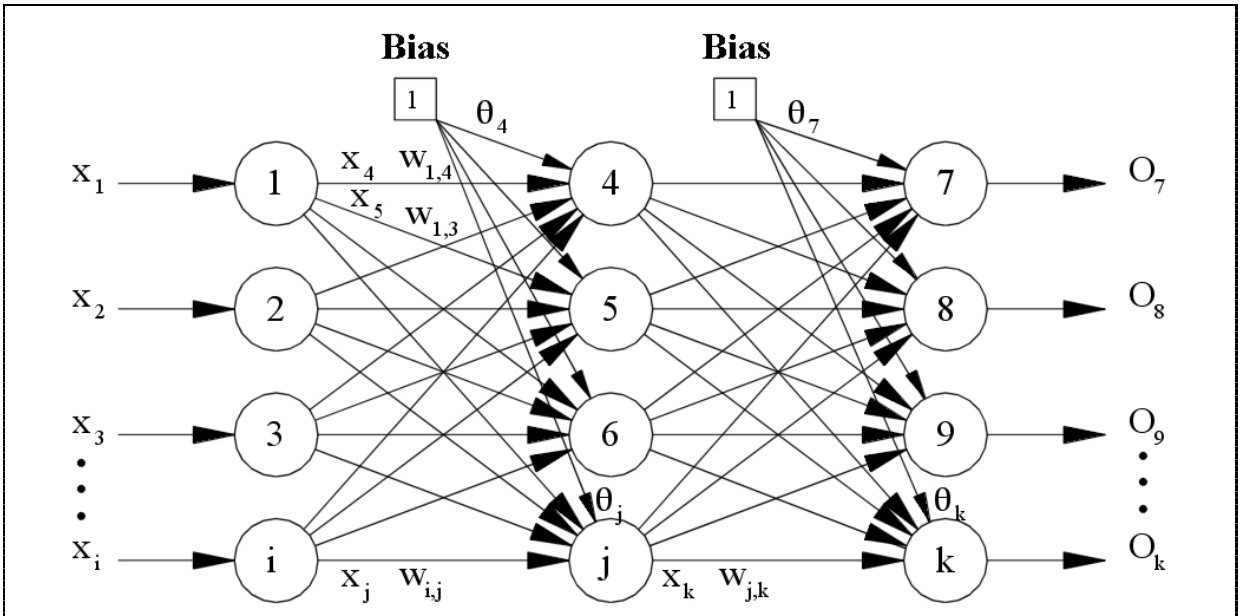

Figure.3. BPNN Architecture

The BPNN is simple, fast, and simple to program. It has no parameters to tune separated from the number of inputs. It is an adaptable technique as it does not need prior information about the network. It is a standard technique that usually performs well. It does not need any specific indications of the features of the function to be learned. Backpropagation is a short term of "backward propagation of errors." It is a standard technique of training ANN [18]. BPNN contains the input, hidden, and output layer, a normal multi-layer network regulated by the complete interconnection within layers but independent of integration in the same cell layer. The process of learning includes forward and backpropagation. The error resulting may be due to the connection weight abnormality

and threshold within the connection layer nodes, thereby measuring the error value of the connecting node and changing it according to the connection weight [19].

Backpropagation refines the network structure by removing weighted links that minimally affect the trained network. It is the technique for adjusting the neural network's weight dependent on the error value acquired in the prior iteration. Proper weight tuning enables to minimize error rates and to make the system dependable by expanding its generalization. The inputs X arrive through the pre-connected path. The input was modeled using weight W. The weights are usually selected in random. The outputs for all neurons from the input layer to the hidden and output layers are calculated. To compute the error in the output, the following condition is used,

Error=Actual output - Desired output

Then switch to the hidden layer from the output layer to adjust the weights, thus reducing the error. This operation is repeated until it obtains the desired outcome. BPNN architecture is depicted in figure.3. It comprises of three layers. When the input was fed with a specific training pattern, the weighted input sum for node j in the hidden layer was determined by,

$$Net_j = \sum w_{i,j} x_j + \Theta_j \quad (1)$$

Condition (1) was to compute the neuron's total input. The Θj expression from a bias node, the weighted value, consistently presented output as one. This bias node was deemed as the "pseudo input" for every neuron present in the output and hidden layers and also utilized for solving the issues identified with circumstances where the input pattern is 0. If an input pattern gets zero value without a bias node, the NN cannot be trained.

To determine whether the neuron must fire, the "Net" term, otherwise called the action potential, was forwarded to a suitable actuation process. The activation function's output value decides the result of neurons and turns into the input to neurons in the following layers associated with it. Hence the differentiable activation function is a prerequisite for the BP algorithm. The sigmoid function was utilized as a standard activation function.

$$O_j = x_k = \frac{1}{1+e^{-Net_j}} \quad (2)$$

Likewise, conditions (1) and (2) were utilized to decide the value of output for node k in the output layer.

Output Layer

If the output node's actual activation value, k, is O_k, and for node k's anticipated target output was t_k, the dissimilarity among the real and the anticipated output was presented by:

$$\Delta_k = t_k - O_k \quad (3)$$

For node k, the error signal in the output layer computed as

$$\delta_k = \Delta_k O_k (1 - O_k) \quad (4)$$

or

$$\delta_k = (t_k - O_k) O_k (1 - O_k)$$

$O_k(1\text{-}O_k)$ is the Sigmoid function derivative. Using delta rule, the weight correlation difference of the node k and j is determined by the j actuation in relation to the k error.

The equations for changing the weight, $w_{j,k}$, among the j and k are:

$$\Delta w_{j,k} = l_r \delta_k x_k \quad (5)$$

$$w_{j,k} = w_{j,k} + \Delta w_{j,k} \quad (6)$$

Where, $\Delta w_{j,k}$ was the modification in the weight among k and j, l_r was the learning rate. l_r is generally a smaller constant, which represents the relevant changes in weight. It must be noticed, in condition (5), the x_k was the value of input for k, and also a similar output value from j. For enhancing the way toward weights updating, a change to condition (5) was

$$\Delta w_{j,k}^n = l_r \delta_k x_k + \Delta w_{j,k}^{(n-1)} \mu \quad (7)$$

The update of weight during the n^{th} epoch was specified through adding the (μ) momentum term, multiplied to $\Delta w_{j,k}$ $(n\text{-}1)^{th}$ epoch.

Hidden Layer

For j, the error signal in this layer computed as

$$\delta_k = (t_k - O_k) O_k \sum (w_{j,k} \delta_k) \quad (8)$$

Likewise, the equation to modify the $w_{i,j}$, within the j and i is

$$\Delta w_{j,k}^n = l_r \delta_j x_j + \Delta w_{j,k}^{(n-1)} \mu \quad (9)$$

$$w_{i,j} = w_{i,j} + \Delta w_{i,j} \quad (10)$$

Global Error

At last, the BP was acquired by presuming, which was appropriate to reduce the output nodes error across the patterns introduced to the NN. The accompanying condition was utilized for computing the error function E, for every pattern:

$$E = {}^1/_2 \sum (\sum (t_k - O_k)^2) \quad (11)$$

Technically, when the NN has been trained adequately, the error function must be a zero value.

3.3. BPNN Algorithm

Allot every network input and output

Initialize each weight with lower numbers randomly, generally among - 1 and 1

repeat

for each pattern in the training set

Present the network pattern

//Propagate the input forward through the network:

for every layer in the network

for each node in the layer

1. Evaluate the input's weighted sum for the node

2. Add the threshold to the sum

3. Evaluate the activation for the node

end

end

//Propagate the errors backward within the network

for each output layer node

evaluate the error signal

end

for every hidden layer

for each layer node

1. Evaluate the error signal of the node

2. Update weight for every node in the network

end

end

//Compute Global Error

Compute the Error Function

end

while ((maximum number of iterations < than determined) AND

(Error Function is > than determined))

From the genetic algorithm, the detected attacks are classified using the BPNN classifier. The experiments performed in the Amazon Cloud Platform with 8 cores and 32 GB of memory. The dataset CICIDS 2017 was used for evaluating the proposed model's performance analysis.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Dataset Description

For detecting the attacks and intrusions, CICIDS 2017 dataset is used. Commonly, various DDoS attack datasets have many constraints like out of relevant data, a redundancy that is variable. This proposed data set has network identical data. The data set was accumulated for five days with different attacks and also normal data. It has the network data with and without attacks, which was approximately the real data of the network [20]. This dataset was uneven, hence this dataset with the duplicating method as it basically affects the training of the proposed model, and then the testing was performed. This research was experimented using Keras on the Tensorflow package on 64-bit Intel Core-i5 processor with 8 GB RAM in Windows 8 system. The algorithms are executed in MATLAB.

Table.1. Class Labels of Dataset with Instances

Class Labels	Number of instances
Bot	1966
BENIGN	2359087
DDoS	41835
DoS Hulk	231072
DoSGoldenEye	10293
DoSslowloris	5796
DoSSlowhttptest	5499
FTP-Patator	7938
Heartbleed	11
Infiltration	36
PortScan	158930
SSH-Patator	5897
Web Attack – Brute Force	1507
Web Attack – SQL Injection	21
Web Attack – XSS	652

4.2. Attack Types

Bot: The attacker use Trojans to breach the security of numerous victim systems, accepting accountability for those systems and set every system in Bot network, which could be utilized and accessed by the intruders remotely.

Benign: Normal traffic action.

DDoS: The attacker uses various systems that works together to perform an attack on one victim system.

DoS Hulk: The intruder use the HULK tool for completing the DoS attack on the web server for making volumes of various and disordered traffics. Likewise, the generated traffics can avoid caching engines and attacks the resource pool of server.

DoSGoldenEye: The intruder uses the GoldenEye tool to implement a DoS attack.

DoSslowloris: The intruder uses the Slow Loris tool to implement the DoS attacks.

DoSSlowhttptest: The intruder utilize the HTTP Get request for outperform the total HTTP links allowed on the server, hindering different clients from incoming and offering the intruders the option to allow several HTTP links with the equivalent server.

FTP-Patator: The intruder uses this to implement the brute force attack for discovering the FTP access details.

Heartbleed Attack: The intruder utilizes the protocol OpenSSL to install malicious data inside OpenSSL memory, allowing the intruder with illegal access to significant data.

Infiltration: The intruder uses the infiltration procedure and softwares to penetrate and acquire unauthorized login to the networked system data.

PortScan: The intruder try to accumulate informations related to the victim system such as the OS details and services running over the packet forwarding with various destinations.

SSH-Patator: The intruder uses this to implement the brute force attack to discover the SSH access details.

Web Attack – Brute Force: The intruder tries to acquire significant information, like PIN and Password using trail-and-error.

Web Attack–SQL Injection: It was a technique of code injection used to attack applications by the data, along with odious SQL proclamation, which were installed inside a part for execution.

Web Attack – XSS: The attacker inject with commonly trusted internet-sites and benign using the web application which redirects to malicious content.

4.3. Performance Metrics

The accuracy was just the performance subset of the model. It is one of the performance metrics for evaluating the classification techniques. The accuracy computation is calculated using the following expression.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

Precision is a positive predictive rate. It is the ratio of properly predicted positive observation to the total predicted positive value. The calculation of precision is calculated using the following expression:

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

The recall is also termed as sensitivity. It is the ratio of properly predicted positive value to the each observation in the actual class. The computation of recall is calculated using the following expression:

$$Recall = \frac{TP}{(TP+FN)} \quad (14)$$

The level of intrusion instances are represented by detection rate. It represents the total proper positive class predictions made as the proportion of all the predictions made. The calculation of the DR is computed using the accompanying condition:

$$DR = \frac{TP}{TP+FN} \quad (15)$$

TP: true positive, FP: false positive, FN: false negative, TN: true negative.

From the dataset, the classes of minority attack having similar behavior and features are combined. Hence combining similar classes, the class, and the predominance ratio of various attack labels seems to be improved. As shown in the table.1, the ratio of Benign class was significant part with 83.34% and where the least attack class was Heartbleed with 0.00039%.

Table.2. Performance Analysis of Detecting Normal Attack (Benign)

Technique	Accuracy	Recall	Precision	Detection Rate
FC-ANN	90.44	89.31	89.25	89.38
NB-RF	92.15	90.12	90.05	90.24
KDBN	94.60	92.58	91.34	93.10
FCM-SVM	96.21	94.24	93.88	94.04
GA-BPNN	97.58	97.98	95.40	96.16

With this significant variation in ratio value, the primary detectors could determine benign. The analysis of the GA-BPNN was computed and compared with various detection methods like FC-ANN (Fuzzy Clustering Artificial Neural Network), NB-RF (Naive Bayes-Random Forest), KDBN (K-Dependency Bayesian Network) and FCM-SVM (Fuzzy C-Means with Support Vector Machine) as represented in table.2. The accuracy and detection rate achieved by the proposed method is 1.3% to 7% higher and 2.1% to 6.7% higher for benign attacks.

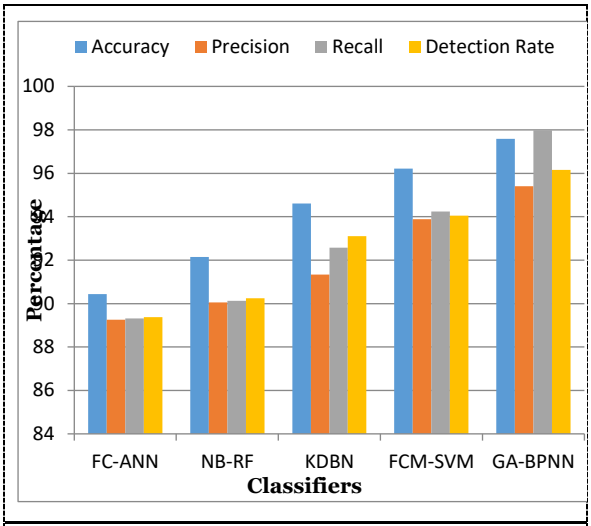

Figure.4. Graphical Plot of Performance Metrics of Benign Detection

The Bot class is called as Botnet ARES. This class contains 1966 instances with the ratio of 0.06%. As shown in the table.3, the proposed method achieved 1.6% to 6.7% more accuracy and 2.3% to 7.4% more detection rate than other techniques for Botnet attacks.

Table.3. Performance Analysis of Detecting Botnet Attack

Technique	Accuracy	Recall	Precision	Detection Rate
FC-ANN	90.25	90.40	89.14	90.05
NB-RF	92.55	91.84	91.01	90.78
KDBN	93.97	93.25	92.42	93.89
FCM-SVM	95.33	94.99	93.98	95.12
GA-BPNN	96.98	96.38	95.16	97.50

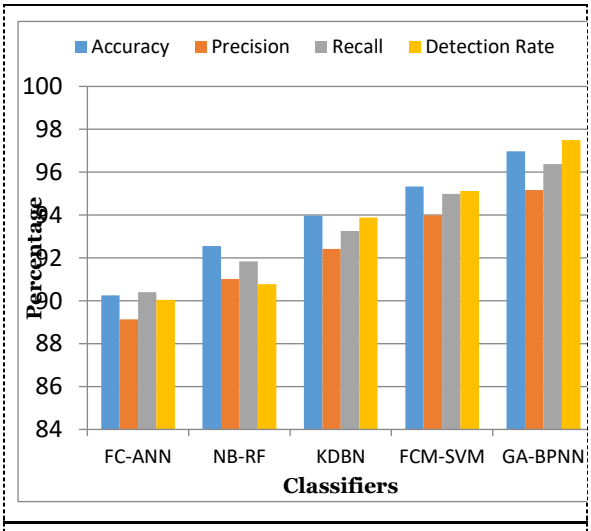

Figure.5. Graphical Plot of Performance Metrics of Botnet Detection

Table.4. Performance Analysis of Detecting Brute Force Attack (FTP & SSH-Patator)

Technique	Accuracy	Recall	Precision	Detection Rate
FC-ANN	90.51	91.04	90.01	91.43
NB-RF	91.34	92.01	90.56	92.24
KDBN	92.97	93.25	91.33	94.00
FCM-SVM	94.64	95.11	93.80	96.18
GA-BPNN	96.89	96.95	95.45	97.89

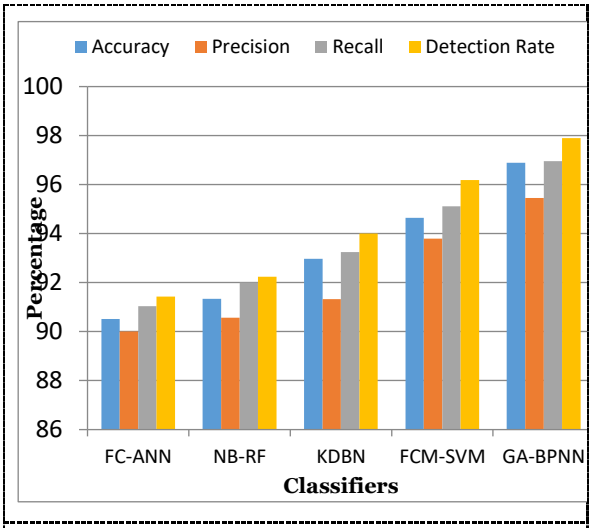

Figure.6. Graphical Plot of Performance Metrics of Brute Force Detection

The SSH-Patator and FTP-Patator classes are united as Brute Force class. Because both the classes have similar features and behavior, by uniting both the classes, a new class was formed with 13835 instances with 0.48% ratio. For this attack type, the proposed method's accuracy is 2.2% to 6.3% higher, and the detection rate is 1.7% and 6.4% higher than other methods, as shown in table.4.

The DoS/DDoS is a new class, which is the combinations of DDoS, DoSHulk, DoSGoldenEye, DoSslowloris, DoSSlowhttptest, and Heartbleed. By combining all these labels, it contains 294506 instances with 10.4% ratio. The GA-BPNN method achieved 0.6% to 4.9% higher accuracy and 1.2% to 6.6% higher detection rate than other techniques for this DoS/DDoS attack detection, as shown in table.5.

Table.5. Performance Analysis of Detecing Dos/DDos Attack ("DDoS, Dos Hulk, DosGoldenEye, DoSslowloris DoSSlowhttptest, & Heartbleed")

Technique	Accuracy	Recall	Precision	Detection Rate
FC-ANN	92.26	92.65	91.50	91.05
NB-RF	93.40	93.97	92.04	93.71
KDBN	95.01	95.55	94.22	95.16
FCM-SVM	96.55	97.11	95.07	96.40
GA-BPNN	97.17	98.10	96.00	97.65

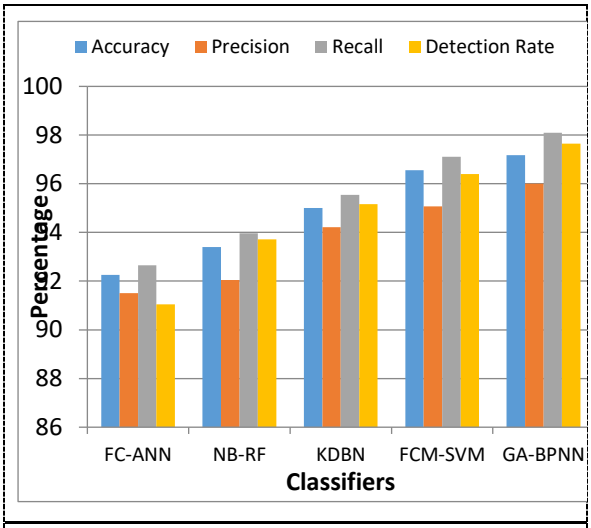

Figure.7. Graphical Plot of Performance Metrics of DoS/DDoS Detection

The performance analysis of the infiltration class and PortScan class are analyzed individually. Both labels are not similar to the features and action of the various classes. Infiltration attack has 36 instances, with a 0.001% ratio, which is the minimum attack ratio of the entire instances. For this infiltration attack detection, the achieved accuracy is 1.5% to 3.8% higher, and the detection rate is 0.7% and 6.8% higher than the other compared methods as represented below in table.6.

Table.6. Performance Analysis of Detecting Infiltration Attack

Technique	Accuracy	Recall	Precision	Detection Rate
FC-ANN	92.35	92.50	91.10	89.58
NB-RF	93.10	93.89	92.44	91.02
KDBN	93.88	94.06	92.65	93.27
FCM-SVM	94.67	94.91	93.16	95.64
GA-BPNN	96.21	96.73	94.98	96.40

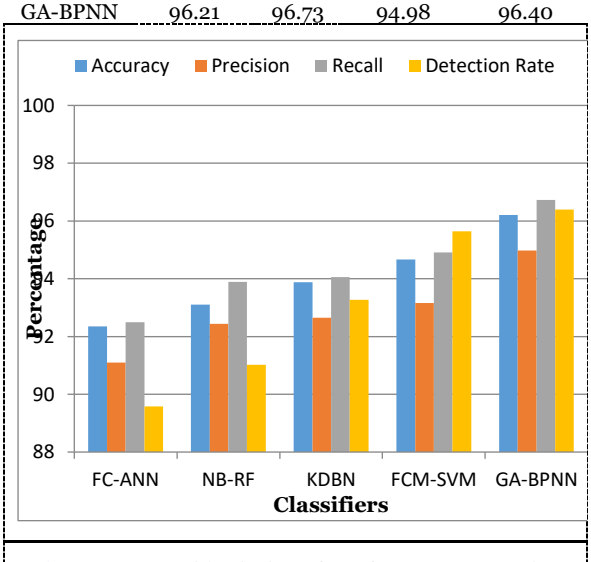

Figure.8. Graphical Plot of Performance Metrics of Infiltration Detection

The PortScan class has 158930 instances, with a 5.61% attack ratio comparing with the entire instances. The PortScan attack detection by GA-BPNN has achieved 1.2% to 7.5% higher accuracy and 1.6% to 6.6% more detection rate than compared techniques, as shown in table.7.

Table.7. Performance Analysis of Detecting PortScan Attack

Technique	Accuracy	Recall	Precision	Detection Rate
FC-ANN	90.12	90.21	89.33	89.32
NB-RF	92.40	93.06	91.10	90.86
KDBN	94.01	93.69	93.45	92.50
FCM-SVM	96.37	95.88	95.48	94.41
GA-BPNN	97.64	97.65	96.57	96.01

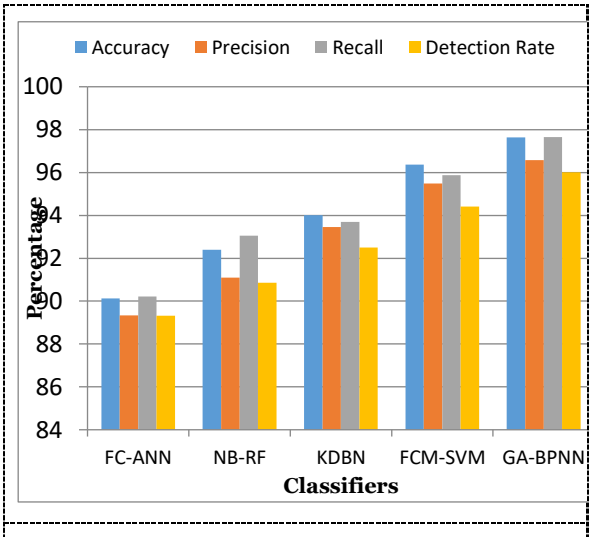

Figure.9. Graphical Plot of Performance Metrics of PortScan Detection

The Web Attack class includes Web Attack-Brute Force, SQL Injection, XSS with 2180 instances and a 0.07% attack ratio. The accuracy achieved by the proposed model for Web attack detection is 1% to 6% higher, and the detection rate is 1.7% to 7.3% higher than the other techniques as represented in table.8.

Table.8. Performance Analysis of Detecting Web Attack ("Web Attack – Brute Force, Web Attack – SQL Injection & Web Attack – XSS")

Technique	Accuracy	Recall	Precision	Detection Rate
FC-ANN	91.85	92.04	90.12	90.51
NB-RF	93.40	93.67	92.54	93.17
KDBN	95.24	96.11	94.10	94.98
FCM-SVM	96.88	96.55	96.04	96.12
GA-BPNN	97.90	97.98	96.76	97.85

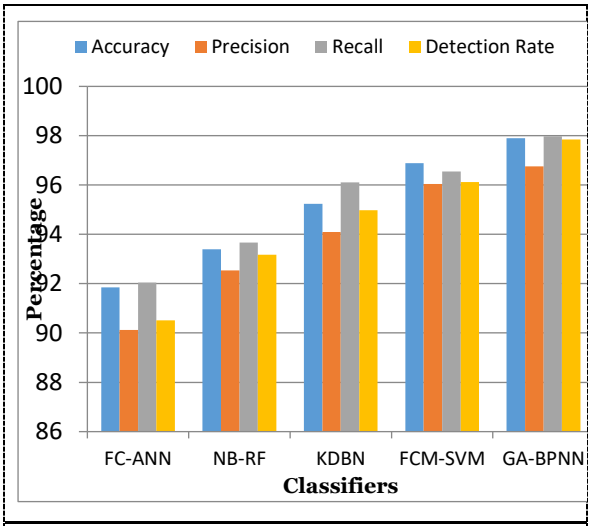

Figure.10. Graphical Plot of Performance Metrics of Web Attack Detection

Overall, the proposed GA-BPNN technique achieved higher accuracy in the detection of Web attacks with 97.90% accuracy and highest detection rate for Brute force attack detection with 97.89%.

V. CONCLUSION AND FUTURE WORK

In this research, an intrusion detection system for the cloud computing environment is proposed. Here in this model, the genetic algorithm and back propagation neural network are used for attack detection and classification. The CIC-IDS 2017 dataset was used for the evaluation of performance analysis. Initially, from the dataset, the data are preprocessed, and by using the genetic algorithm, the attack was detected. The detected attacks are classified using the back propagation neural network classifier for identifying the types of attacks. The performance analysis was executed, and the results are obtained and compared with the existing machine learning-based classifiers like FC-ANN, NB-RF, KDBN, and FCM-SVM techniques. The proposed GA-BPNN model outperforms all these classifying techniques in every performance metric, like accuracy, precision, recall, and detection rate. Overall, from the performance analysis, the best classification accuracy is achieved for Web attack detection with 97.90%, and the best detection rate is achieved for Brute force attack detection with 97.89%. In the future, the proposed research can be experimented as an IDS model for the Internet of Things (IoT) platform to detect and classify attacks.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

None.

REFERENCES

[1] Rajkumar Buyya, James Broberg, and Andrzej Goscinski. (2011). Cloud Computing Principles and Paradigms. John Wiley & Sons Inc.

[2] Chirag Modi et al. (2013). A survey on security issues and solutions at different layers of Cloud computing. Journal of Supercomputing. Springer. Vol.63, pp.561-592.

[3] Loubna Dali et al. (2015). A Survey of Intrusion Detection System. 2nd World Symposium on Web Applications and Networking (WSWAN). IEEE. pp.1-6.

[4] Manisha Rani and Gagandeep. (2019). A Review of Intrusion Detection System in Cloud Computing. International Conference on Sustainable Computing in Science, Technology & Management. pp.770-776.

[5] Ahmed Patel, Mona Taghavi, Kaveh Bakhtiyari, and Joaquim Celestino Junior. (2013). An intrusion detection and prevention system in cloud computing: A systematic review. Journal of Network and Computer Applications. Elsevier. Vol.36, pp.25-41.

[6] Ahmed Shawish and Maria Salama. (2014). Cloud Computing: Paradigms and Technologies. Inter-cooperative Collective Intelligence: Techniques and Applications, Studies in Computational Intelligence. Springer. Vol.495, pp.39-67.

[7] Mostapha Derfouf and Mohsine Eleuldj. (2019). Implementations of Intrusion Detection Architectures in Cloud Computing. International Conference of Cloud Computing Technologies and Applications. Springer. pp.100-124.

[8] Yasir Mehmood, Umme Habiba, Muhammad Awais Shibli, and Rahat Masood. (2013). Intrusion Detection System in Cloud Computing: Challenges and Opportunities. 2nd National Conference on Information Assurance. IEEE. pp.59-66.

[9] S N Dhage et al. (2011). Intrusion Detection System in Cloud Computing Environment. International Conference and Workshop on Emerging Trends in Technology. pp.235-239.

[10] Nureni Ayofe Azeez et al. (2019). Intrusion Detection and Prevention Systems: An Updated Review. Data Management, Analytics and Innovation. Vol.1042, pp.685-696.

[11] Nurudeen Mahmud Ibrahim and Anazida Zainal. (2020). A Distributed Intrusion Detection Scheme for Cloud Computing. International Journal of Distributed Systems and Technologies. Vol. 11, No.10, pp.68-82.

[12] Partha Ghosh, Sumit Biswas, Shivam Shakti, and Santanu Phadikar. (2020). An Improved Intrusion Detection System to Preserve Security in Cloud Environment. International Journal of Information Security and Privacy. Vol.14, No.1, pp.67-80.

[13] Ahmad Shokouh Saljoughi, Mehrdad Mehvarz, and Hamid Mirvaziri. (2017). Attacks and Intrusion Detection in Cloud Computing Using Neural Networks and Particle Swarm Optimization Algorithms. Emerging Science Journal. Vol. 1, No. 4, pp.179-191.

[14] Bahram Hajimirzaei and Nima Jafari Navimipour. (2019). Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. ICT Express. Vol.5, No.1, pp.56-59.

[15] Mohamed Idhammad, Karim Afdel, and Mustapha Belouch. (2018). Distributed Intrusion Detection System for Cloud Environments based on Data Mining techniques. Procedia Computer Science. Elsevier. Vol.127, pp.35-41.

[16] Aws Naser Jaber and Shafiq Ul Rehman. (2020). FCM–SVM based intrusion detection system for cloud computing environment. Cluster Computing Springer. https://doi.org/10.1007/s10586-020-03082-6.

[17] Hongsheng Yin et al. (2019). Intrusion Detection Classification Model on an Improved k-Dependence Bayesian Network, Special Section on Innovation and Application of Intelligent Processing ff Data, Information and Knowledge as Resources in Edge Computing. IEEE Access. Vol.7, pp.157555-157563.

[18] Muhammad Salman Taj, Syed Irfan Ullah, Dr.Abdus Salam, and Wajid Ullah Khan. (2020). Enhancing Anomaly Based Intrusion Detection Techniques for Virtualization in Cloud Computing Using Machine Learning. International Journal of Computer Science and Information Security. Vol.18, No.5, pp.68-78.

[19] Roshani Gaidhane, C. Vaidya, and M. Raghuwanshi. (2014). Intrusion Detection and Attack Classification using Back-propagation Neural Network, International Journal of Engineering Research & Technology. Vol.3, No.3, pp.1112-1115.

[20] Razan Abdulhammed et al. (2019). Features Dimensionality Reduction Approaches for Machine Learning Based Network Intrusion Detection. Electronics. Vol.8, No.332, 2019, pp.1-27.

Cite this article as: Suresh A N. A Hybrid Genetic-Neuro Algorithm for Cloud Intrusion Detection System. J. Comput. Sci. Intell. Technol. 2020; 1(2): 15–25. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 2, pages: 26–33, September 2020

A Review on Prostate Cancer Detection using Deep Learning Techniques

[1]Narmatha C, [2]Surendra Prasad M

[1]*Department of Computer Science, Faculty of Computers and Information Technology, University of Tabuk, Saudi Arabia.*
[2]*Annapoorana Medical College and Hospitals. Salem, Tamil Nadu, India*

[**]***Corresponding Author: surendar1745@gmail.com***

Received: *10.06.2020,*
Revised: *10.07.2020,*
Accepted: *19.08.2020,*
Published: *31.09.2020*

DOI:
10.53409/mnaa.jcsit20201204

Abstract: The second most diagnosed disease of men throughout the world is Prostate cancer (PCa). 28% of cancers in men result in the prostate, making PCa and its identification an essential focus in cancer research. Hence, developing effective diagnostic methods for PCa is very significant and has critical medical effect. These methods could improve the advantages of treatment and enhance the patients' survival chance. Imaging plays a significant role in the identification of PCa. Prostate segmentation and classification is a difficult process, and the difficulties fundamentally vary with one imaging methodology then onto the next. For segmentation and classification, deep learning algorithms, specifically convolutional networks, have quickly become an optional technique for medical image analysis. In this survey, various types of imaging modalities utilized for diagnosing PCa is reviewed and researches made on the detection of PCa is analyzed. Most of the researches are done in machine learning based and deep learning based techniques. Based on the results obtained from the analysis of these researches, deep learning based techniques plays a significant and promising part in detecting PCa. Most of the techniques are based on computer aided detection (CAD) systems, which follows preprocessing, segmentation, feature extraction, and classification processes, which yield efficient results in detecting PCa. As a conclusion from the analysis of some recent works, deep learning based techniques are adequate for the detection of PCa.

Keywords: *Prostate Cancer, CAD system, MRI, Ultrasound, Segmentation, Classification*

I. INTRODUCTION

PCa is the second common diagnosed disease and one of the critical diseases to cause mortality among men around the world. The growth of overall PCa is expected to reach around 2.3 million new cases and 740000 deaths by 2040 normally because of the aging and growth of the population. In view of the GLOBOCAN database 2018, the incidence and mortality rate is represented in figure 1 and 2 [1].

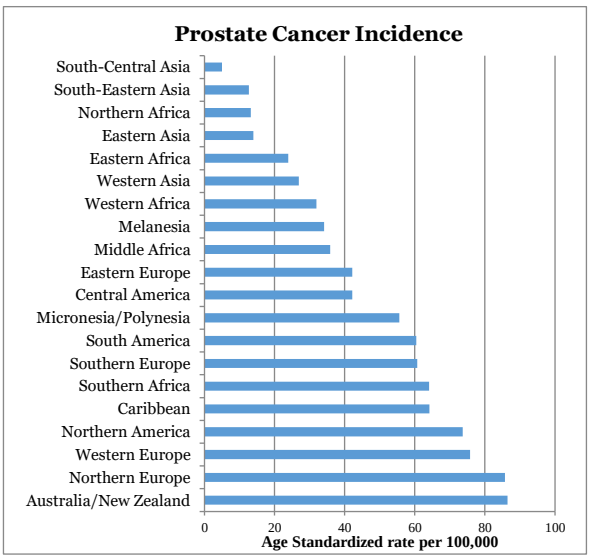

Figure.1. Prostate Cancer Incidence Chart of 2018

In 2018, the number of new cases recorded all over the world was 1,276,106, which covered 7.1% of cancer patients excluding other cancers

globally. And 358,989 deaths were occurred in 2018 globally due to PCa, which covers 3.8% of mortality. It is one of frequently diagnosing cancer around men in more than one-half (105 of 185) of the world nations.

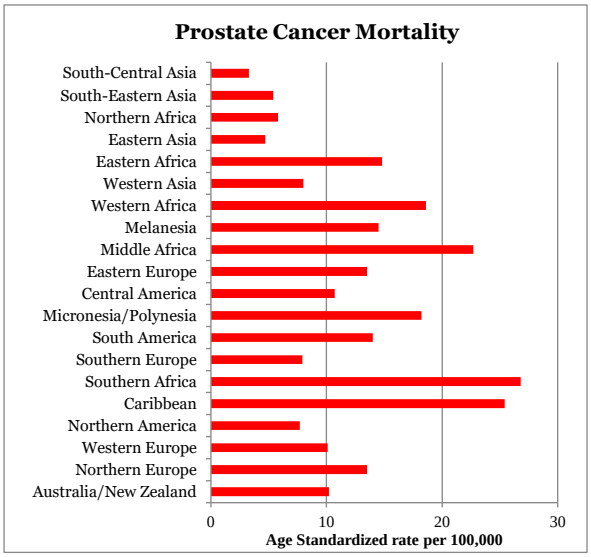

Figure.2. Prostate Cancer Mortality Chart of 2018

The higher incidence rate and death rates worldwide are shown in figure.2. PCa has conventionally been diagnosed by digital rectal exam (DRE) and prostate-specific antigen (PSA) blood test, trailed by transrectal ultrasound (TRUS) or multi-parametric MRI (mpMRI) guided biopsy. As PCa is a multiform disease, extending from small, slow, low-grade tumors, to large, intense, dangerous tumors, the essential objective for urologists during baseline assessment of prostate disease is, subsequent to determining the presence of the tumor, to assess local and distant cancer expansion, and its grade by staging [2]. However, Transrectal Ultrasound (TRUS) directed biopsy was as yet the principle methodology for diagnosing PCa [3], it was not suggested because of its intrusiveness [4].

Rapid technological advances in the course of the most recent years have empowered the standard utilization of imaging prostate for the clinical direction of PCa. Imaging models like multiparametric-MRI (mpMRI), multiparametric ultrasound (mpUS), and Positron Emission Tomography (PET) nuclear imaging were currently being utilized for every aspect of PCa diagnoses and localizations, staging, focal and whole-gland therapy, recurrence monitoring, and active surveillance. MRI has demonstrated as a major compelling imaging method for PCa, enhancing localizations and direction of biopsies, particularly for anterior cancers [5]. T1 & T2 stage infection that was bound to the prostate was effectively perceptible on MRI, while additional prostatic diseases (T3 and T4) were difficult for visualization.

Table.1. Advantages and Disadvantages of Imaging Modalities

Imaging Models	Clinical Usage	Advantages	Drawbacks	Prospective
Ultrasound-based	Early diagnosis and detection	Widely available, Office-based, real-time imaging, inexpensive	Limited tissue contrast among cancerous and benign tissue	mpUS-based method (RTE, CEUS) may enhance contrast
mpMRI-based	Early diagnosis and recurrence, active surveillance, staging, metastatic involvement	Best tissue contrast for detection of medically significant Prostate Cancer	High-cost due to in-bore time, lack of real-time imaging, needs advanced training	Alternative in-bore options with real-time imaging being advanced
mpMRI-ultrasound fusion-based	Early diagnosis and detection, active surveillance	Combines multimodality data, Office-based	Most costly, needs either fusion-device specific training or ample experience to execute cognitive fusion, registration errors during MRI-ultrasound fusion	Gaining popularity worldwide, but additional enhancements to mini-mize registration errors required
PET-based	Staging, recurrence, metastatic spread	Offers ancillary details for tumor staging, characterization and metastatic involvement	High-cost, technological (i.e. attenuate +-ion correction) and/or medical challenges (i.e. radiation effects)	Development of specific radionuclides is an ongoing endeavor

As a result of disease's variability, analysts and clinicians in recent years have begun utilizing multiparametric MRI (mpMRI), consolidating anatomic and functional imaging by many other successions to give a progressively comprehensive image. The mpMRI diagnosis commonly comprise diffusion weighted imaging (DWI) for cellularity, T2-w for anatomy, and for vascularity, dynamic contrast enhanced MRI (DCE-MRI) [6]. Various clinical analyses have evaluated the efficiency of mpMRI for PCa, and have decided it was the finest non-invasive alternation for detecting cancer. The mpMRI has can alter the ideal models on PCa detection and classification of risk [7]. The mpMRI of the prostate was basically some operative imaging

types utilized to complement normal anatomical T1 and T2-weighted images. The operative arrangements of options are DWI and DCE, Imaging plays an inexorably significant part in the initial recognition and the direction of PCa. The key imaging models, mpMRI, mpUS, PET, and MRI–US fusion are utilized in the localization and diagnosing PCa. Significance was based on conditions of tumors biologically that maintain the utilization of particular imaging methods [9]. PCa diagnosis was generally depended on various methods as digital rectal examination, PSA, TRUS, MRI, and transrectal biopsy. Specifically, the DWI-MRI method enables for acquiring images with contrast contingent upon the tissue's microscopic water molecules mobility, testing the microscopic structures. Additionally from DWI, image was conceivable for evaluating the Apparent comprising the apparent diffusion co-efficient (ADC) maps calculation [8].

Diffusion Coefficient (ADC) of water utilizing diverse diffusion model, as "Monoexponential", "Bi-exponential", "Kurtosis", "Gamma conveyance" and "Stretched exponential", each one of them dependent on various speculations on the tissue's microenvironment water mobility [10] [12]. Multi-parametric MRI could present detailed representation of prostate lesions and tissues. The cancer could be recognized earlier any essential invasive methodologies like needle biopsy, at the risk of harm or periprostatic nerves irritation, bladder neck and prostate. In any case, the prostate tissue cancer on MRI could likewise be hard to detect, with frequently uncertain results between the clinicians [13-14].

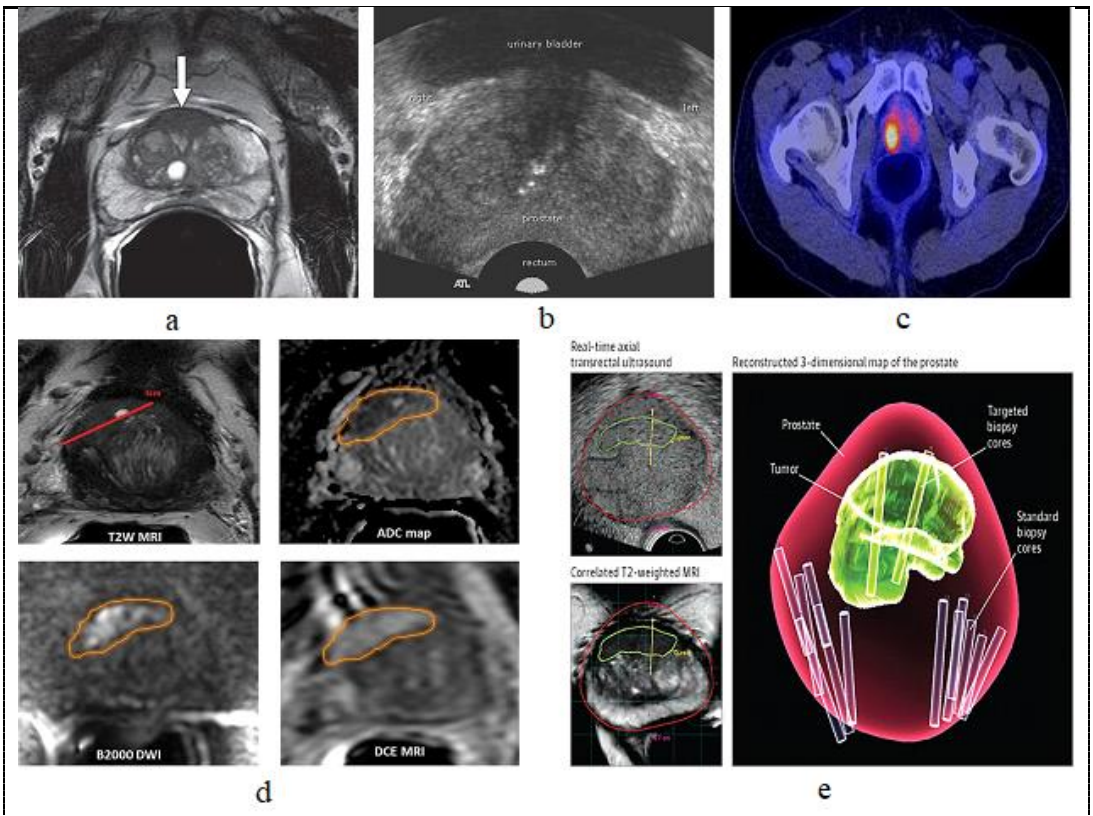

Figure.3. Diagnosis of Prostate Cancer using, a) MRI, b) Ultrasound, c) PET, d) mp-MRI, and e) MRI-US fusion [10]

II. RELATED WORK BASED ON SEGMENTATION AND CLASSIFICATION

2.1. Segmentation based Analysis

Segmenting PCa was a difficult assignment, and the difficulties fundamentally vary starting with one imaging methodology then onto the next. Micro-calcifications, speckle, low contrast, and imaging artifacts such as shadow presents critical difficulties to segment exact prostate in TRUS. Be that as it may, in MRI, superior soft tissues contrast emphasizes huge variations in size, texture, and shape data within the prostate. Interestingly, poor soft tissue differentiates among the prostate and encompassing tissue in CT image makes difficulty in segmenting precise prostate segmentation [15] [16].

Precise and solid prostate gland segmentation utilizing MRI has the significance in diagnosing and medication of prostate cancer. Although several automatic segmentation techniques depended on deep learning and machine learning has been proposed, the performance of segmentation has opportunity to get better because of the huge variation in image appearance, interference, and anisotropic spatial resolutions. A 3D adversarial pyramid anisotropic convolutional deep neural network for segmenting MRI prostate was used in [17]. The FCNN–based network structures for segmenting MRI PCa image was introduced in [18] and analyzed different structures of alternate associations along with the deep network size and recommend 8 diverse FCNN–based deep 2D network structure for automated MRI segmentations of prostate. In [19], an ACM (Active Contour Model) was trained and utilized for segmenting the PCa. Then, after injection, features were extricated from the dynamic MRIs at various times and changed them into RIC curve. Then, discriminative features were chosen by SFFS (Sequential Forward Floating Selection) and FDR (Fisher's Discrimination Ration).

By using the Salp Swarm Optimization Algorithm-based Rider Neural Network (SSA-RideNN) and Color Space (CS) transformation the segmentation and classification was done in [20]. Detection of the prostate disease types from CT images of abdomen utilizing texture analysis was analyzed in [21]. Texture features were extricated from the images segmented utilizing an evolving transformation called Sequency based Mapped Real Transform (SMRT). This feature sets were obtained by changing and block size and sub-image size. Every feature set was optimized utilizing a Genetic Algorithm (GA). The finest feature set was chosen dependent on accuracy of classification. A fully automated segmentation algorithm for T2-w endorectal prostate MRI was proposed in [22] and assessed its accuracy inside various ROI utilizing a set of complementary error measurements. A neural-fuzzy technique for automatic region segmentation in TRUS images was proposed in [23].

In [26] a non-invasive system for the early recognition of PCa from DW-MRI was analyzed. The prostate was localized and segmented dependent on the new level-set. To maintain sequence, the computed ADC values were refined and normalized utilizing the image model of Generalized Gauss-Markov Random Field (GGMRF). The multi-layer deeply supervised deconvolution network (DSDN) that executes end-to-end training for automatic segmentation of MRI. More layers deeply supervised were added to supervise the hidden layers performances [27].

2.2. Classification based Analysis

The convolutional neural network is one of the feed-forward neural networks, which gives advantages in classification of image undertakings as per practical experience. Among many deep learning algorithms, the CNN is an algorithm executes better in the classification process. The CNN as an image classification technique that produced some diagnosis classification references as discussed in [28]. A DBN-DS-based multi-classifier prediction model for pathologic stage prostate cancer was used for classification in [29]. The classifiers were made by learning information as indicated by classifier. To assess the DBN-DS-based multi-classifier, the whole dataset was separated into the training and testing sets. In [30], a deep learning system, which depends on the 3D CNN, to extricate the spatial-temporal features consistently from the images of sequential CEUS was used to detect tumor. For obtaining the CEUS image sequence's temporal patterns, a 3rd dimension was combined in the CNN model.

A complete automatic CNN-based CAD system for initial diagnoses of PCa was analyzed in [31]. Prostate delineation depended on the level set which utilized 3 sorts of attributes for enhanced performances. These attributes were shape priors, spatial voxel relationships, and appearance. The integration of the shape priors improved the delineation accuracies as most prostates had analogous structure. These three types of features were combined utilizing a non-negative matrix factorization (NMF) method. An imaging-based methodology in the prediction of 3-years biochemical recurrence (BCR) through a novel SVM classifier was analyzed in [32]. AdaBoost-based Ensemble Learning was used in [33] for supporting automatic Gleason grading of prostate adenocarcinoma (PRCD). In [34], an automatic Clinically Significant Prostate Cancer recognition system where every process was optimized together in an end-to-end trainable DNN was analyzed. Three sorts of loss functionalities like classification, overlap, and inconsistency losses were utilized to optimize every parameter of the CNN and TDN.

Table.2. Analysis of Performance Evaluations

Author	Year	Performance	Technique	Modality	Dataset	Results	
Zhiyu Liu et al. [13]	2019	Segmentation and Classification	Mask R-CNN and DNN	T2W-MRI	PROSTATEx and I2CVB	AUC	0.882 0.912
Haozhe Jia et al. [16]	2019	Segmentation	3D APA-net	T2W-MRI	Promise12 ASPS13 Hybrid	DSC	0.906 0.893 0.901
Tahereh H et al. [17]	2019	Segmentation	FCNN	MRI	Promise12	DSC	0.874
Chuan-Yu C et al. [18]	2015	Feature based Classification	SVM	MRI	Own	ACC	94.74
Shashidhar B G [19]	2019	Classification	Improved RideNN	Prostate histopathological images	Own	ACC	89.66
Manju B et al. [20]	2015	Feature based Classification	Genetic algorithm and KNN	CT	Own	ACC	94.30
Maysam Shahedi et al. [21]	2017	Segmentation	Fully automated segmentation algorithm	T2W-MRI	Own	DSC	82
Islam Reda et al. [22]	2016	Segmentation and Classification	NMF-Autoencoder	DW-MRI	Own	ACC	97.6
Ying Liu and Xiaomei An [23]	2017	Classification	CNN	DWI-MRI	ImageNet	ACC	78.15
Dong Ji et al. [24]	2018	Segmentation	DSDN	MRI	Dataset from Medical Images Computing and Computer Assisted Interventions	DSC	90.15
Jae Kwon Kim et al. [25]	2018	Classification	DBN-DS	TRUS-Pathology	KPCR	ACC AUC	81.27 0.777
Yujie Feng et al. [26]	2018	Feature Extraction and Classification	3D-CNN	CEUS-Ultrasound	CEUS	ACC	90.18
Islam Reda et al. [27]	2018	Feature Extraction and Classification	ADCs-based CNN	DWI-MRI	Own	ACC	97.60
Yu-Dong Zhang et al. [28]	2016	Classification	SVM	mp-MRI	Own	ACC	92.2
Chao-Hui H and Emarene M K [29]	2016	Classification	AdaBoost-based Ensemble Learning	PRAD histopathological image	TCGA	ACC	97.8
Zhiwei Wang et al. [30]	2017	Classification	TDN and CNN	mp-MRI	PROSTATEx	AUC	0.962
Yanan Shao et al. [31]	2020	Classification	GAN	Ultrasound	CRCEO, JH, PCC, PMCC, and UVA	AUC	0.934
R.J.G. van Sloun et al. [32]	2019	Segmentation and Classification	FCNN	TRUS	Own	ACC DSC	98 0.96
Adeel Ahmed M et al. [33]	2020	Classification	CNN	MRI	Dataset by Harvard University	AUC	1.00

III. DISCUSSION

Imaging plays a significant role in the prostate cancer detection. Imaging models like mpMRI, mpUS, PET, and MRI–US fusion imaging are utilized in the localization and diagnosis of PCa. For detecting the prostate cancer segmentation and classification are the two fundamental processes to carry out the performance analysis. Prostate segmentation is a difficult process, and the difficulties altogether vary from one modality to another. Micro-calcifications, Low contrast, speckle, and imaging artifacts such as shadow poses are the major difficulties to precisely segment the prostate in medical images. The present CAD system analyzes images from different imaging modalities like US and MRI for detecting and localizing PCa and also to assess its size and stage. The CAD system is effective and utilized for recognizing PCa in the central gland (CG), peripheral zone (PZ), and transition zone from MRI T2-weighted. In this study, the detection of PCa based on MRI imaging was considered and reviewed. Further, the CAD processes for PCa like pre-processing, segmentation, feature extraction and classification is discussed on the plot of MRI based performance.

3.1. Imaging Datasets

The Cancer Imaging Archive (TCIA) has 9 prostate disease data sets accessible from http://www. cancerimagingarchive.net/:

- Prostate-MRI dataset with (26 cases) of prostate MRIs.
- Prostate-Diagnosis project with (92 cases), PCa T1-and T2-w MRIs.
- NaF Prostate with (9 cases) was an assortment of F-18 NaF CT/PET images in patients with PCa.
- The Prostate-3T venture (64 cases) gave Prostate transversal T2-weighted MRIs information to TCIA as a feature of an ISBI challenges competition in 2013.
- The QIN PROSTATE dataset with (22 cases) of the Quantitative Imaging Network (QIN) includes multi-parametric MRIs gathered for the task of detection and staging of prostate cancer.
- The TCGA-PRAD dataset with (14 cases), additionally referenced in the Genomics part of this analysis, and likewise has imaging data (Pathology, CT, PET, and MRI).
- The Prostate Fused-MRI-Pathology dataset with (28 cases) was a fusion of histopathology slides and MRI images.
- The dataset PROSTATEx Challenge with (346 cases) was a prostate review set of MR analysis.
- The dataset QIN-PROSTATE-Repeatability with (15 cases) was a data set with multi-parametric PCa MRI implemented in a test-retest setting, enabling to assess the MRI-based analysis repeatability in the PCa.

3.2. Segmentation

The process of segmentation comprises of representing the MRI's prostate boundaries and specific significance for concentrating the posterior process on organ of interest. The process was to segment the ROIs or prostate from the dataset of 2D or 3D prostate data. Different types of datasets are publicly available in three domains: imaging data, clinical data, and genomics data.

This segmentation process could be performed in automatic, self-automatic, or manually. Automated segmentation is a difficult process due to the patient's action or move during image scanning, the inter patients difference of prostate's appearance and shape, inhomogeneities, and noise, disintegration of prostate boundaries as a result of occlusion, and neighboring tissue's comparable intensities.

However the prostate boundary provides the locations of prostate gland volumes while segmenting manually, it needs additional time and was observers based. Accordingly, various segmentation techniques are designed to consider those issues and accurately plot tissues of prostate like deformable and statistical based techniques. The deformable models (DMs) were generally utilized to segment the prostate from DWI. Statistical based techniques were utilized to plot the prostate from MRI methods. The conclusive outcomes, however promising, indicated that the prostate segmentation issue is as yet uncertain.

3.3. Classification

The last process of a CAD framework includes training and testing with features extricated from labels and images. Many measurements can be utilized so as to survey the performance of a classifier. The main metric utilized was the accuracy which was calculated as the true identification ratio to the samples count. Though, based upon the technique utilized in the CAD system, this metric could be exceptionally biased through a higher count of true negative instances that would increase the accuracy overrating of the real classifier performance.

The most general statistical computations are specificity and sensitivity which provide a full outline of the classifier's efficiency. Sensitivity was additionally known as true positive value and was equivalent to the samples of true positives ratio beyond true positives included with samples of false negatives. Specificity was likewise known as rate of true negatives and was equivalent to the proportion of the samples of true negatives beyond the true negatives included with samples of false positives. These measurements could be utilized to figure the ROC-(Receiver Operating Characteristic) curves. A statistics obtained from ROC analysis was the AUC- Area Under the Curve that relates to the area under ROC and was an estimation utilized to make comparison among techniques. At last, besides the overlap measures, the Dice's coefficients were normally calculated to assess the lesions localization accuracies.

IV. ANALYSIS

This segmentation and classification graphical representations are plotted using the results observed from different techniques in the survey. This comparison chart is plotted based on the results obtained from using only MRI images like T2W, DWI, and mp-MRI. Because most of the researches are based on MRI imaging modalities and MRI based research produced best results in terms of detecting prostate cancer efficiently and effectively. In both segmentation and classification the deep learning based techniques performed well with the results. As shown in the figure the CNN is the top most choice of the

researchers for segmenting and classifying prostate cancer. The researchers integrate the deep learning techniques with other classifiers to design hybrid techniques for efficient results. The CNN has the advantages of automatically detects the significant features without any manual supervision and it is computationally efficient. Comparing CNN with other techniques like SVM and different machine learning techniques, for example, CNN is suitable for large datasets where SVM is not and SVM are margin classifier and support different kernels to perform the classification. The deep learning techniques performed better on ultrasound, CT and histopathological images too.

V. CONCLUSION AND FUTURE WORK

This survey presents the overview of imaging techniques used for PCa and the review of segmentation and classification techniques. It was difficult to point out which imaging technique or which classification method is suitable for the PCa detection. Based on the literature review and the analysis of the performance of existing and recent works made on prostate cancer detection, this survey is concluded. Dataset plays an important role in the segmentation and classification process. Different types of datasets are publicly available in three domains: imaging data, clinical data, and genomics data. In the process of detection or classification of PCa, an efficient and accurate algorithm is necessary. Most of the detection techniques are based on CAD- computer aided detection system. In which, preprocessing, segmentation, feature extraction, and classification are the major processes to be carried out for efficient results. Based on this efficient result, this survey is representing the deep learning based classifiers particularly Convolution Neural Network (CNN) was very often used by many researchers for the classification of PCa detection. This type of deep learning classifier was suitable for both segmentation and classification. In future, hybrid techniques based on deep learning and machine learning could be efficient for the task of PCa detection. A review of PCa, imaging modalities and the techniques used for segmentation and classification of the PCa was analyzed and surveyed in this paper. This survey will be useful for researchers intended to do a research on detecting and classifying PCa in future.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

None.

REFERENCES

[1] Mary B B. Culp, Isabelle S, Jason A. E, Freddie B, and Ahmedin J. (2019). Recent Global Patterns in Prostate Cancer Incidence and Mortality Rates. *Eur Urol. Elsevier*. pp.1-15.

[2] Jean-L D. (2019). Diagnosis of prostate cancer. *Asian J Urol. Science Direct*. Vol.6, pp.129-136.

[3] Raman P S, Savita G, Rajendra A. (2017). Segmentation of prostate contours for automated diagnosis using ultrasound images: A survey. *J Comput Sci. Elsevier*. pp.1-31.

[4] Ahmed S et al. (2018). Computer-Aided Diagnosis of Prostate Cancer on Diffusion Weighted Imaging: A Technical Review. *IEEE International Conference on Imaging Systems and Techniques (IST)*. pp.1-6.

[5] Hüseyin C D and John W D. (2018). Multiparametric magnetic resonance imaging: Overview of the technique, clinical applications in prostate biopsy and future directions. *Turk J Urol*. Vol.44, No.2, pp.93-102.

[6] Tristan B. (2015). What is multiparametric-MRI of the prostate and why do we need it? *Imag Med*. Vol.7, No.2, pp.13-17.

[7] Bejoy A and Madhu S N. (2019) On Computer-Aided Diagnosis of Prostate Cancer from MRI using Machine Intelligence Techniques. *IEEE International Conference on Electrical, Computer and Communication Technologies*. pp.1-8.

[8] Latrach A, Trigui R, Chenini H, Sellemi L, and Ben H A. (2018). Comparison Study for Computer Assisted Detection and Diagnosis 'CAD' systems Dedicated to Prostate Cancer Detection Using MRImp Modalities. *International Conference on Advanced Technologies for Signal and Image Processing. IEEE*. 2018, pp.1-6.

[9] Saradwata S and Sudipta D. (2016). A Review of Imaging Methods for Prostate Cancer Detection. *Biomed Eng Comput Biol*. Vol.7, No.S1, pp.1-15.

[10] Andreas M and Thomas F. Clinical value of multi-parametric ultrasound and MRI/US fusion-guided biopsy for prostate cancer detection and visualization. *Canon Medical Systems*.

[11] Andrea B et al. (2018). A review on the role of water Diffusion modeling in Magnetic Resonance Imaging of Prostate cancer. *IEEE Workshop on Complexity in Engineering (COMPENG)*. pp. 1-5.

[12] Alexander S L and Arvid L. (2019). An overview of deep learning in medical imaging focusing on MRI. *Z Med Phys. Elsevier*. Vol.29, pp.102-127.

[13] Zhiyu L et al. (2019). A Two-Stage Approach for Automated Prostate Lesion Detection and Classification with Mask R-CNN and Weakly Supervised Deep Neural Network. *Artif Intell Radiat Ther. Springer*. pp.43-51.

[14] Guillaume L, Robert M, Jordi F, Joan C. V, Paul M. W, and Fabrice M. (2015). Computer-Aided Detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: A review. *Comput Biol Med. Elsevier*. Vol.60, pp-1-54.

[15] Shijun W, Karen B, Baris T, Peter C, and RonaldM S. (2014). Computer Aided-Diagnosis of Prostate Cancer on Multiparametric MRI: A Technical Review of Current Research. *BioMed Research International. Hindawi*. Vol.2014, pp.1-11.

[16] Haozhe J, Yong X, Yang S, Donghao Z, Heng H, Yanning Z, and Weidong C. (2020). 3D APA-Net: 3D Adversarial Pyramid Anisotropic Convolutional Network for Prostate Segmentation in MR Images. *IEEE T Med Imaging*. Vol.39, No.2, pp.447-457.

[17] Tahereh H, Len H, and Kevin H S. (2016). Convolutional Neural Networks for Prostate Magnetic Resonance Image Segmentation. *IEEE Access*. Vol.4, pp.1-12.

[18] Chuan-Yu C, Hui-Ya H, and Yuh-Shyan T. (2015). Prostate Cancer Detection in Dynamic MRIs. *IEEE International Conference on Digital Signal Processing (DSP)*. pp.1279-1282.

[19] Shashidhar B. G, Kshama V. K, and Veena V. D. Prostate Cancer Detection using Histopathology Images and Classification using Improved RideNN. *Biomed Eng-App Bas C*. Vol.31, No.6, pp.1-14.

[20] Manju B, K.Meenakshy, and R. Gopikakumari. (2015). Prostate Disease Diagnosis from CT images using GA optimized SMRT based Texture Features. International Conference on Information and Communication Technologies. *Procedia Computer Science*. Vol.46, pp.1692-1699.

[21] Maysam Shahedi et al. (2017). Accuracy Validation of an Automated Method for Prostate Segmentation in Magnetic Resonance Imaging. *J digit imaging*. Vol.30, pp.782-795.

[22] Islam Reda et al. (2016). A New NMF-Autoencoder based CAD System for Early Diagnosis of Prostate Cancer. *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. pp.1237-1240.

[23] Ying L and Xiaomei A. (2017). A Classification Model for the Prostate Cancer Based on Deep Learning. International Congress on Image and Signal Processing. *BioMed Eng Inform*. IEEE. pp.1-6.

[24] Dong J, Jun Y, Toru K, Liangfeng X, and Shu Z. (2018). Automatic Prostate Segmentation on MR Images with Deeply Supervised Network. *International Conference on Control. Decision and Information Technologies. IEEE*. pp.309-314.

[25] Jae K K et al. (2018). A Deep Belief Network and Dempster-Shafer-Based Multiclassifier for the Pathology Stage of Prostate Cancer. *J Healthc Eng. Hindawi*. Vol.2018, pp.1-8.

[26] Yujie F, Fan Y, Xichuan Z, Yanli G, Fang T, Fengbo R, Jishun G, and Shuiwang J. (2019). A Deep Learning Approach for Targeted Contrast-Enhanced Ultrasound Based Prostate Cancer Detection. *IEEE/ACM T Comput Biol Bioinform*. Vol. 16, No. 6, pp. 1794-1801.

[27] Islam R et al. (2018). A Novel ADCs-Based CNN Classification System for Precise Diagnosis of Prostate Cancer. *International Conference on Pattern Recognition (ICPR). IEEE*. pp.3923-3928.

[28] Yu-Dong Z et al. (2016). An imaging-based approach predicts clinical outcomes in prostate cancer through a novel support vector machine classification. *Oncotarget*. Vol.7, No.47, pp.78140-78151.

[29] Chao-Hui H and Emarene M K. (2016). Automated Classification for Pathological Prostate Images using AdaBoost-based Ensemble Learning. *IEEE Symposium Series on Computational Intelligence (SSCI)*. pp. 1-4.

[30] Zhiwei W, Chaoyue L, Danpeng C, Liang W, Xin Y, and K.T. Tim Cheng. (2018). Automated Detection of Clinically Significant Prostate Cancer in mp-MRI Images based on an End-to-End Deep Neural Network. *IEEE T Med Imaging*. vol. 37, no. 5, pp. 1127-1139.

[31] Yanan S, Jane W, Brian W, and Septimiu E S. (2020). Improving Prostate Cancer (PCa) Classification Performance by Using Three-Player Minimax Game to Reduce Data Source Heterogeneity. *IEEE T Med Imaging*. pp.1-11.

[32] R.J.G. van Sloun et al. (2018). Zonal segmentation in transrectal ultrasound images of the prostate through deep learning. *IEEE International Ultrasonics Symposium (IUS)*. pp.1-4.

[33] Adeel A A et al. (2020). Detecting prostate cancer using deep learning convolution neural network with transfer learning approach. *Cogn Neurodynamics*. pp.1-11.

Cite this article as: Narmatha C and Surendra Prasad M, A Review on Prostate Cancer Detection using Deep Learning Techniques. J. Comput. Sci. Intell. Technol. 2020; 1(2): 26–33. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 2, pages: 34–41, September 2020

Classification of Alzheimer's disease from MRI Images using CNN based Pre-trained VGG-19 Model

Manimurugan S

[1]*Department of Computer Engineering, Faculty of Computers and Information Technology, University of Tabuk, Saudi Arabia.*

[**]***Corresponding Author: semanimurugan@gmail.com, mmurugan@ut.edu.sa***

Received: *15.06.2020,*
Revised: *21.07.2020,*
Accepted: *19.08.2020,*
Published: *31.09.2020*

DOI:
10.53409/mnaa.jcsit20201205

Abstract: The most typical form of dementia is Alzheimer's disease (AD), which can permanently cause memory cell damage. Since Alzheimer's is a progressive disease, various problems increase over time. For this reason, diagnosing the disease early is very necessary. The primary objective of this work is to use Magnetic Resonance Imaging to diagnose Alzheimer's disease. A Convolution Neural Network (CNN) model, VGG-19, is proposed in this research. VGG-19 is typically a pre-trained deep transfer learning model that is used here to use MRI images to identify Alzheimer's disease. The primary processes performed in this research are preprocessing, feature extraction, and classification. From the OASIS database, the dataset used in this work is obtained. A total of 373 MRI images are used for evaluation. For training, 80% of data and testing, 20% of data are used in this model for performance analysis. The proposed VGG-19 model is evaluated and compared with existing deep learning-based other CNN models like AlexNet, GoogLeNet, and VGG-16. The performance metrics like accuracy, recall, precision, specificity, and f-measure are evaluated to estimate the performance analysis of the model. For every validation, both training and testing results are evaluated and compared. The model has obtained 96.04% training accuracy and 95.82% testing accuracy, which was 1.9% to 5.3% higher than the AlexNet and GoogLeNet models in performance. However, this model is proposed to classify the AD from brain MRI images and obtain better validation results.

Keywords: Alzheimer's disease, Dementia, CNN, VGG-19, OASIS database, MRI.

I. INTRODUCTION

Dementia is a disorder in which memory, actions, perception, and the ability to perform daily tasks deteriorate. The dementia affects the older people mostly and it is not a general part of aging. The most common type of dementia is AD and might lead to 60-70% of cases. Around 50 million people have dementia worldwide, according to the WHO, with about 60 percent living in low- and middle-income countries. There are about 10 million new cases each year. The average percentage at a given time of the general population aged 60 and over with dementia is between 5-8 percent. It is predicted that the total number of people with dementia will surpass 82 million in 2030 and 152 million in 2050. The loss of tissues and death of nerve cells in the brain is caused by AD, leads to human memory loss and having a bad effect on the output of everyday life activities like writing, communicating, and reading. Patients with AD can also have trouble in recognizing their members of family. Patients in the mild cognitive stage act aggressive, but people in the final AD state suffer from cardiac arrest and death-related respiratory dysfunction [1]. For dementia scaling, the GDS (global deterioration scale) was widely utilized. This scale additionally divides dementia into 7 stages, considering stages 1-3 as stages of pre-dementia, and considering stages 4-7 as dementia.

1. No cognitive decline
2. Age-Associated Memory Impairment
3. Mild Cognitive Impairment
4. Mild Dementia
5. Moderate Dementia
6. Moderately Severe Dementia
7. Severe Dementia

In the field of detecting AD, neuroimaging techniques help the evaluation of brain changes and are promising. MRI plays a vital role in neuroimaging techniques in offering both the anatomical and physiological functions of the body. In diagnosing AD, MRI is the most frequently used technique where any defects present in the brain can be detected [1]. MRI-based images provide multi-modal information that is appropriate for clinical purposes of the brain's function and structure.

Many Computer-Aided Diagnostic Systems (CADS) have been developed by researchers to precisely detect and classify the extracted features of AD. Otherwise, additional time and effort by specialists is needed to process the extracted features. Researchers have recently developed deep learning techniques/models to explicitly extricate features from medical images. Deep learning models have obtained significant performance in medical images, like X-ray, MRI, CT, microscopy, and mammography. The key emphasis of these models or approaches was on binary classification, which indicates whether or not the patient is suffering from AD [2].

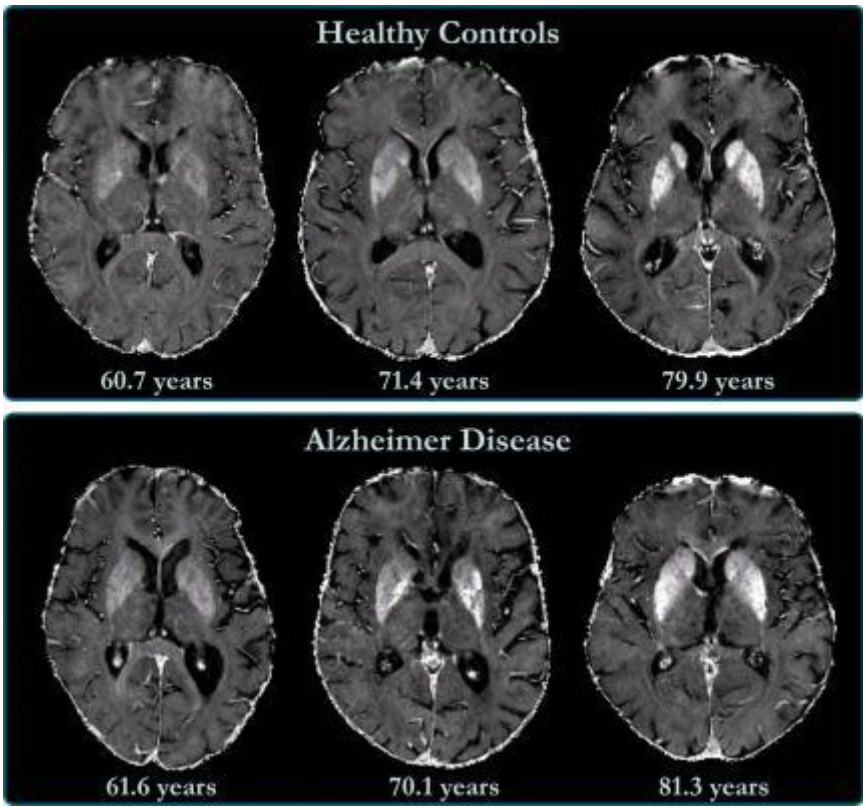

Figure.1. Brain MRI view of difference between Healthy and Alzheimer Disease

In this research, a pre-trained deep convolution neural network model VGG-19 with transfer learning is proposed to analyze and classify AD. This research's main objective is to detect AD based on a CNN model VGG-19 using MRI images. The MRI images collected from the OASIS database are used to evaluate the proposed model in detecting AD. The assessment of performance analysis is based on accuracy, precision, recall, specificity, and F-measure.

II. RELATED WORK

Braulio Solano Rojas et al. proposed a three-dimensional Densenet-121 architecture for AD identification using MRI images. The images were gathered and preprocessed from the ADNI database. Some good performance results have been obtained by this prediction model, and still, to produce more than average results, the model can be improved. A downside may be the reduction of the dataset used for training and testing [3].

Atif Mehmood et al. developed the Siamese-CNN (SCNN) for AD classification. The model was totally similar but joins two parallel layers of modified VGG16, called "Siamese." In this work, the OASIS dataset was used. The initial stage was preprocessing of data and augmentation, the next was extracting

features, and finally, classification. To obtain the maximum features on a limited dataset, the VGG16 and an additional Conv layer were added to the model. A CNN based VGG-16 model was used for classification. Four dementia stage categories have been classified and better performance has been achieved [4].

Taranjit Kaur and Tapan Kumar Gandhi used a pre-trained model DCNN-VGG-16 with transfer learning (TL) for the identification of abnormalities in brain MRI. The features extricated by CNN were highly dependent on the training dataset's size. From pre-trained models, weights of the 'conv' layer were utilized in TL, and just the final layers were trained with data from the new classes. In the context of the elimination of the ROI delineation's preprocessing stages, feature extraction, and selection, the performances were better to the current conventional methods recorded for the classification of brain image [5].

Xiaoling Lu et al. used the MobileNet, a CNN technique for the classification of AD using MRI images. The MRI images were initially preprocessed and using the convolution layer portion, the bottleneck features were extracted. Then to create a new network structure, the newly built completely connected layer was connected. TL was used to load weights on new networks and train networks for pre-training. The MRI images were eventually identified and compared with the model VGG-16. In this work, MobileNet is observed to be less complex and better than VGG-16 [6].

Muhammed Yildirim and Ahmet Cinar proposed a hybrid classification model based on CNN for classifying AD using MRI images. For this hybrid model, the ResNet50 model was used as a basis, and the output was individually tested using different CNN models such as AlexNet, ResNet50, DenseNet201, and VGG16. The last five layers of Resnet50 were removed in the hybrid model, ten new layers were added in place of those layers removed, and the number of layers increased from 177 to 182. The efficiency of the hybrid model has achieved greater accuracy than others at 90 percent compared to all models [7].

Ronghui Ju et al. presented a model for early recognition of AD using deep learning with brain networks. The brain network was built by evaluating the brain region's functional connectivity using data from resting-state functional MRI (R-fMRI). To detect an early AD state, normal aging from mild cognitive impairment, a targeted autoencoder network was implemented. The proposed technique effectively revealed discriminative brain network features and produced a steady AD detection classifier [8].

Marcia Hon and Naimul Mefraz Khan proposed the TL-based AD detection technique using VGG-16 and Inception v4 models from MRI images. The two models were both trained and evaluated and used the entropy-based approach to select the training dataset. In detecting AD, the inception model outperformed the VGG-16 model with an accuracy of 96.25% [9].

Xin Hong et al., using the LSTM technique, proposed a model for predicting AD. A special type of recurrent neural network that could connect previous data to the current task is long short-term memory. The LSTM network with completely linked layers and activation layers was therefore constructed in order to encode the temporal relationship among features and the next stage of AD [10]. F.J. Martinez Murcia et al. proposed a new research data analysis of AD using the deep convolutional autoencoder technique. Using regression and classification analysis, the distribution of the extracted features in various combinations were analyzed and visualized. The impact on the brain of each co-ordinate of the autoencoder manifold was evaluated. The assessment was performed by using the ADNI MRI dataset and achieved an accuracy of over 80 percent [11].

Mingxia Liu et al. proposed a weakly-supervised densely linked neural network (wiseDNN) for brain disease prognosis using baseline MRI data and incomplete clinical scores [12]. Naimul Mefraz Khan et al. proposed a TL-based technique for AD diagnosis from MRI. With layer-wise tuning, the network was fine-tuned, where only a group of pre-defined layers was trained on MRI images. The entropy approach was utilized to decrease the training dataset size and achieved 95.19 percent accuracy [13].

III. PROPOSED METHODOLOGY

CNN is extensively recognized for its capability to execute highly accurate medical image classification in deep learning. However, compared to traditional machine learning techniques, the most significant benefit of CNNs is that it does not need manual extraction of features because CNNs can automatically extract the features and then classifies the AD stages. The principle of TL is to train a pre-trained CNN to use a smaller dataset from a different problem to learn new image representations. TL is a framework for machine learning design; it is not a model or

technique for machine learning. For pre-trained models, this design approach usually applies. These pre-trained models are based on Deep Evolution Neural Networks. In deep learning, this technique involves the initial training of a CNN using large-scale training datasets for a classification issue. Figure 2 demonstrates the architecture of the proposed VGG-19 model.

As shown in Figure 3, VGG-CNN has six main structures, mostly comprised of multi-connected convolutional layers and fully-connected layers. 3*3 is the convolutional kernel dimension, and 224*224*3 is the input size. In general, every layer is concentrated at 16~19. As a preprocessing model, VGG-19 was used. It has been increased in network complexity compared with conventional CNNs. It uses an alternating structure, which is well over a single convolution, with several convolution and non-linear activation layers. The layer structure can extract better features of image, using Maxpooling for downsampling, and adjusting the ReLU as the activation function, selecting the largest value in the image region as the region's pooled value. The downsampling layer is primarily utilized to enhance the image anti-distortion capability of the network while maintaining the main sample features and minimizing the number of parameters [14].

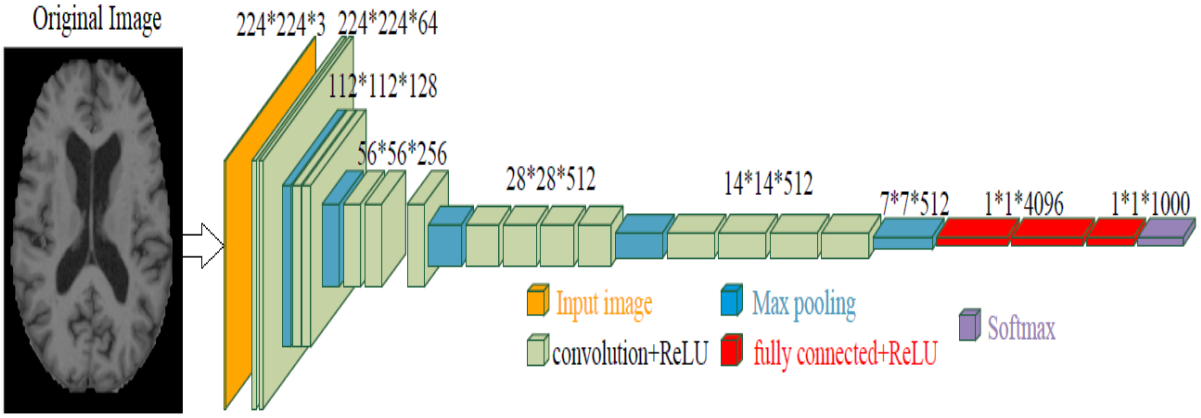

Figure.2. Proposed VGG-19 Architecture Model

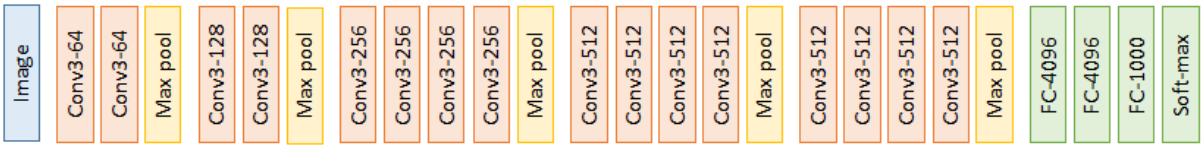

Figure.3. Layers of VGG-19

For the downsampling layer, the expression is presented in equation (1). Among them, $down\left(X_j^{(n-1)}\right)$ is the maximum pooling sampling function. τ_j^n is the coefficient according to the j[th] feature map of the n[th] layer, and $f(\tau_j^n down(\tau_j^n down(X_j^{(n-1)}) + b_j^{(n)})$ is the ReLU activation function.

$$X_{p_j}^{(n)} = f\left(\tau_j^n down\left(X_j^{(n-1)}\right) + b_j^{(n)}\right) \quad (1)$$

The activation function of ReLU is expressed as follows,

$$f(x) = \begin{cases} 0, & x \le 0 \\ x, & x > 0 \end{cases} \quad (2)$$

The activation function of the softmax layer is expressed as follows,

$$f(x_j) = \frac{e^{x_j}}{\sum e^{x_i}} \quad (3)$$

In these above equations, *f(x)* is the activation function, and *x* is the activation function input.

The initial two layers are convolutional layers with 3x3 filters, and 64 filters are used in the first two layers, so it finishes with a volume of 224x224x64 since the same convolutions have been used (height and width are the same). So this (CONV64)x2 reflects that there are two Conv layers with 64 filters in the model.

- The filters are always 3x3 with a stride of 1, and the same convolutions are always applied.
- A pooling layer is then used, which lowers a volume's height and width: it goes from 224x224x64 down to 112x112x64.
- Then it uses a few more layers of Conv. 128 filters are used here, and the same convolutions are used; 112x112x128 would be a new dimension.
- Then, a pooling layer is added so that the new dimension will be 56×56×128.
- Two conv layers with 256 filters are added next.

- A pooling layer
- A few more conv layers with 512 filters
- A pooling layer
- A few more conv layers with 512 filters
- A pooling layer
- In the end, the model has its final 7x7x512 in a Fully-connected layer (FC) with 4096 units and one of 1000 classes in a softmax output.

With a larger collection of data over a smaller one, CNN typically performs well. In those CNN applications where the dataset is not large transfer learning (TL) may be useful. TL's concept utilizes the learned model from large datasets such as ImageNet for applications with comparatively smaller datasets. This removes the need for a large dataset and decreases the lengthy training time as needed when generated from scratch by the deep learning algorithm. TL is a deep learning technique that uses a model trained as a starting point for a single assignment to train a model for a similar assignment. Fine-tuning a network with TL is usually much faster and easier than training a network from scratch. It enables training models using similar small labeled data by leveraging standard models that have already been trained on large datasets. It is possible to dramatically decrease the training time and computing resources. With TL, the model does not need to be trained for as many epochs (a complete training period on the entire dataset) as a new model.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed CNN's performance analysis with the VGG-19 model is assessed using the dataset in this section. The model is evaluated using parameters such as accuracy, recall, precision, specificity, and f-measure. Also, a comparative analysis is conducted for the validation of the model proposed. The output is compared to other current deep learning models used for classification, such as AlexNet, GoogLeNet, and VGG-16. On the MATLAB 2019a Simulink toolbox, all the experiments are simulated and analyzed. The dataset is split into 80 percent for training and 20 percent for testing.

4.1. Dataset Description

The proposed performance review of CNN with the VGG-19 model is assessed in this section using the dataset. Using parameters such as accuracy, recall, precision, specificity, and f-measure, the model is evaluated. A comparative analysis is also performed for the validation of the proposed model. The performance is compared with other existing deep learning models, such as AlexNet, GoogLeNet, and VGG-16, used for classification. All the experiments are simulated and analyzed on the MATLAB 2019a Simulink toolbox. The dataset is split into 80 percent for training and 20 percent for testing.

In this research, the OASIS dataset is used for the evaluation. The Open Access Series of Imaging Studies (OASIS) aims to make neuroimaging data sets of the brain freely available to the scientific community. The dataset consists of 150 subjects aged 60 to 96 years in a longitudinal collection. For a total of 373 imaging sessions, every subject was diagnosed for two or more visits, separated by at least one year. 3 or 4 individual T1-weighted MRI scans acquired in single scan sessions are included for each subject. Across the study, 72 of the subjects were labelled as non-demented. At the time of their initial visits, 64 of the subjects included were labelled as demented and remained so for subsequent scans, including 51 people with mild to moderate AD. Based on the CNN model, data preprocessing was the significant part of obtaining effective and precise results for those algorithms. The image size of the OASIS dataset was 256*256, but an image size of 224*224 is needed for the proposed CNN model.

4.2. Performance Metrics

In this research, the proposed CNN with the VGG-19 model is proposed and compared with other deep learning CNN models such as AlexNet, GoogLeNet, and VGG-16. To estimate the performance analysis of the model, the performance metrics such as precision, accuracy, recall, specificity, and f-measure are evaluated. Both training and testing outcomes are measured and compared with every validation. The primary objective of this research is to identify AD from MRI brain images, which can be helpful in assessing the patient's abnormality. This model's outcome may be dependent on the result detected as normal or abnormal. True positive (TP), true negative (TN), false positive (FN), and false negative (FN) are correctly evaluated to predict this model's outcome.

TP: It refers the total correct predictions in abnormal cases.

FP: It refers the total incorrect predictions in abnormal cases.

TN: It refers the total correct predictions in normal cases.

FN: It refers the total incorrect predictions in normal cases.

Accuracy is the estimate of the performance subset by the model. It is the primary performance metric used to measure the classification process efficiency. Where both the positive and negative groups are equally significant, it is generally used for evaluation. It is calculated using the equation below,

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Precision is a predictive value that is positive. The preciseness of the classification model is computed by using it. It is the estimate of the correctly predicted positive observation's cumulative predictive positive value. The lower precision value reflects a large number of false positives, which affects the classification model. The measure of precision can be computed using the equation below,

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Sensitivity is also known as recall. It is the proportion of correctly predicted positive observation of the overall positive predictive value. The lower recall value reflects a large number of false-negative values, which affects the classification model. The recall estimation can be calculated using the equation below,

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

According to this model, specificity is the prediction that healthy subjects do not have an abnormality. It is the percentage of subjects with no injury/trauma that is tested as abnormal. The specificity estimation can be calculated using the equation below,

$$Specificity = \frac{TN}{TN+FP} \quad (7)$$

The F-measure estimates the performance accuracy and is specified as the weighted harmonic mean of the precision and the recall. The accuracy does not take into account how the data is distributed. The f-measure is then used to manage the distribution problem with accuracy. The f-measure estimation can be calculated using the following equation,

$$F-measure = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (8)$$

Table.1. Performance Analysis of Training Evaluation

Model	Accuracy	Precision	Recall	Specificity	F-measure
AlexNet	94.13	93.81	93.64	95.25	93.72
GoogLeNet	95.66	94.22	93.80	95.90	95.39
VGG-16	95.84	94.90	94.92	96.02	95.56
VGG-19	96.04	95.96	95.75	96.18	95.81

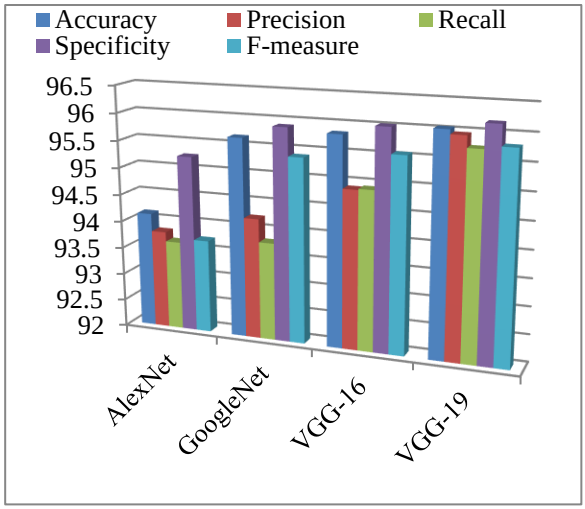

Figure.4. Performance Analysis of Training Data

Table.2. Performance Analysis of Testing Evaluation

Model	Accuracy	Precision	Recall	Specificity	F-measure
AlexNet	89.45	88.74	88.23	90.07	89.19
GoogLeNet	92.90	91.55	91.06	93.12	91.09
VGG-16	94.91	93.32	93.35	95.31	94.24
VGG-19	94.82	93.24	93.13	95.14	94.10

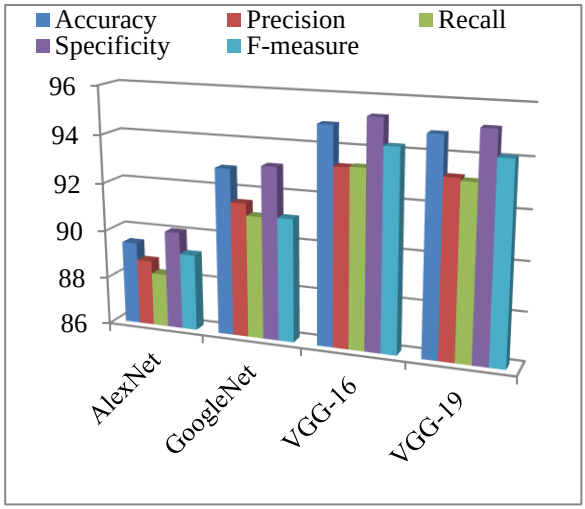

Figure.5. Performance Analysis of Testing Data

As shown in table.1 and 2, the proposed VGG-19 model achieved better performance in training and testing for classifying the ACL tear MRI images. The model obtained 96.04% training accuracy and 95.82% testing accuracy, which is 1.9% to 5.3% increased performance than the other existing compared models. The AlexNet, GoogLeNet, and VGG-16 achieved 94.13%, 95.66%, 95.84% in training and 89.45%, 92.90%, 94.91% in testing

performance respectively. The VGG-16 and VGG-19 model achieved better performance in testing compared to the performance. The performance was the same as accuracy for other parametric evaluations like precision, recall, specificity, and f-measure.

Compared with both training and testing results, the AlexNet model achieved the least performance in every parameter. The VGG-16 model has achieved close results to the proposed model and better accuracy on testing performance. The graphical chart of the comparison is plotted in figure.4 and 5.

V. CONCLUSION AND FUTURE WORK

In this research, a CNN based VGG-19 model was proposed for classifying the AD using MRI images. The proposed model was executed into four stages; Data preprocessing is the initial stage. After preprocessing, the next stage is to extract the image's features and then deliver the extracted features into the classifier to train. The model after training is finally tested. The dataset used in this work was collected from the OASIS database. It is an open-access database for the brain MRI dataset. A total of 373 MRI images were used for evaluation. For training, 80% of data (298 images) and testing, 20% of data (75 images) were used in this model for performance analysis. The proposed VGG-19 model was evaluated and compared with existing deep learning-based other CNN models like AlexNet, GoogLeNet, and VGG-16. The performance metrics like accuracy, precision, recall, specificity, and F-measure were evaluated to estimate the performance analysis of the model. For every validation, both training and testing results were evaluated and compared. The model has obtained 96.04% training accuracy and 95.82% testing accuracy, which was 1.9% to 5.3% higher than the AlexNet and GoogLeNet models in performance. However, this model was proposed to classify the AD from brain MRI images and obtained better validation results. In future, the proposed model can be used to find any abnormalities present in various body organs by using different datasets like a brain tumor, heart disease, lung cancer, etc. The performance can be further increased by implementing a new deep transfer learning model with better classification performance.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The authors confirm that the data supporting the findings of this research are available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

None.

REFERENCES

[1] Mamata L and Rashmi P. (2018). A Survey on Classification Methods of Brain MRI for Alzheimer's disease. Int J Eng Res Technol. Vol.7, No.5, pp.339-348.

[2] Suhad Al-Shoukry, Taha H. R, and Nasrin M. M. (2020). Alzheimer's Diseases Detection using Deep Learning Algorithms: A Mini-Review. IEEE Access. Vol.8, pp.77131-77141, DOI: 10.1109/ACCESS.2020.2989396.

[3] Solano-R B., Villalón-F R., Marín-R G. (2020). Alzheimer's Disease Early Detection Using a Low-Cost Three-Dimensional Densenet-121 Architecture. In: Jmaiel M., Mokhtari M., Abdulrazak B., Aloulou H., Kallel S. (eds) The Impact of Digital Technologies on Public Health in Developed and Developing Countries. ICOST 2020. Lecture Notes in Computer Science, Vol.12157, pp.3-15. Springer, Cham. https://doi.org/10.1007/978-3-030-51517-11.

[4] Atif M, Muazzam M, Muzaffar B, and Yang S. (2020). A Deep Siamese Convolution Neural Network for Multi-Class Classification of Alzheimer Disease. Brain Sciences. Vol.10, No.84, pp.1-15.

[5] Taranjit K and Tapan K G. (2019). Automated brain image classification based on VGG-16 and transfer learning. 2019 International Conference on Information Technology (ICIT), pp.94-98, DOI: 10.1109/ICIT48102.2019.00023.

[6] Xiaoling L, Haifeng W, and Yu Z. (2019). Classification of Alzheimer's disease in MobileNet. Journal of Physics: Conference Series. Vol.1345, 042012, DOI:10.1088/1742-6596/1345/4/042012.

[7] Muhammed Y and Ahmet C. (2020). Classification of Alzheimer's disease MRI Images with CNN Based Hybrid Method. Ingénierie des Systèmes d'Information. Vol.25, No.4, pp.413-418. https://doi.org/10.18280/isi.250402.

[8] Ronghui J, Chenhui H, Pan Z, and Quanzheng L. (2019). Early Diagnosis of Alzheimer's disease based on Resting-State Brain Networks and Deep

Learning. IEEE/ACM T Comput Biol Bioinform, Vol.16, No.1, pp.244-257, DOI: 10.1109/TCBB.2017.2776910.

[9] Marcia H and Naimul M K. (2017). Towards Alzheimer's disease Classification through Transfer Learning. 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp.1166-1169. DOI: 10.1109/BIBM.2017.8217822.

[10] Xin H, Rongjie L, Chenhui Y, Nianyin Z, Chunting C, Jin G, and Jane Y. (2019). Predicting Alzheimer's Disease Using LSTM, IEEE Access, Vol.7, pp.80893-80901, DOI: 10.1109/ACCESS.2019.2919385.

[11] F.J. Martinez-Murcia, A. Ortiz, J.M. Gorriz, J. Ramirez, and D. Castillo-Barnes. (2020). Studying the Manifold Structure of Alzheimer's disease: A Deep Learning Approach Using Convolutional Autoencoders, IEEE J Biomed Health, Vol.24, No.1, pp.17-26, DOI: 10.1109/JBHI.2019.2914970.

[12] Mingxia L, Jun Z, Chunfeng L, and Dinggang S. (2020). Weakly Supervised Deep Learning for Brain Disease Prognosis Using MRI and Incomplete Clinical Scores. IEEE T Cybernetics. Vol.50, No.7, pp.3381-3392, DOI: 10.1109/TCYB.2019.2904186.

[13] Naimul M K, Nabila A, and Marcia H. (2019). Transfer Learning with Intelligent Training Data Selection for Prediction of Alzheimer's disease. IEEE Access. Vol.7, pp.72726-72735, DOI: 10.1109/ACCESS.2019.2920448.

[14] Jian X, Jia W, Shaozhong C, and Bilong L. (2020). Application of a Novel and Improved VGG-19 Network in the Detection of Workers Wearing Masks, Journal of Physics: Conference Series Vol.1518, 012041, DOI: 10.1088/1742-6596/1518/1/012041.

Cite this article as: Manimurugan S., Classification of Alzheimer's disease from MRI Images using CNN based Pre-trained VGG-19 Model. J. Comput. Sci. Intell. Technol. 2020; 1(2): 34–41. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 3, pages: 01 –08, December 2020

A New Neural Network-Based Intrusion Detection System for Detecting Malicious Nodes in WSNs

[1]Narmatha C

[1]Department of Computer Science, Faculty of Computers and Information Technology, University of Tabuk, Tabuk City, Saudi Arabia.

***Corresponding Author: cmnarmath@gmail.com and narmatha@ut.edu.sa**

Received: *02.09.2020,*
Revised: *05.12.2020,*
Accepted: *15.12.2020,*
Published: *22.12.2020*

DOI:
10.53409/mnaa.jcsit20201301

Abstract: The Wireless Sensor Networks (WSNs) are vulnerable to numerous security hazards that could affect the entire network performance, which could lead to catastrophic problems such as a denial of service attacks (DoS). The WSNs cannot protect these types of attacks by key management protocols, authentication protocols, and protected routing. A solution to this issue is the intrusion detection system (IDS). It evaluates the network with adequate data obtained and detects the sensor node(s) abnormal behavior. For this work, it is proposed to use the intrusion detection system (IDS), which recognizes automated attacks by WSNs. This IDS uses an improved LEACH protocol cluster-based architecture designed to reduce the energy consumption of the sensor nodes. In combination with the Multilayer Perceptron Neural Network, which includes the Feed Forward Neutral Network (FFNN) and the Backpropagation Neural Network (BPNN), IDS is based on fuzzy rule-set anomaly and abuse detection based learning methods based on the fugitive logic sensor to monitor hello, wormhole and SYBIL attacks.

Keywords: Magnetic Resonance Imaging, Brain tumor, and Deep Belief Network.

I. INTRODUCTION

Wireless sensor networks (WSNs) are current technology, and researchers have attracted a lot of attention. Usually, the low power and low-cost environment of the WSN includes a large number of sensors that are arbitrarily distributed or reassigned manually over the target location. Due to its potential features and applications such as healthcare, monitoring, domestic uses, surveying systems, and disaster management [1], wireless sensor networks are becoming a powerful and familiar technology. There are poor communication, calculation, and energy capacities for wireless sensor nodes. Broadcast messages are a useful and essential prototype in wireless sensor networks, which permit multiple users to efficiently combine and distribute Message Packages across their network to obtain their data of interest. Figure 1 shows an example diagram for WSN [2].

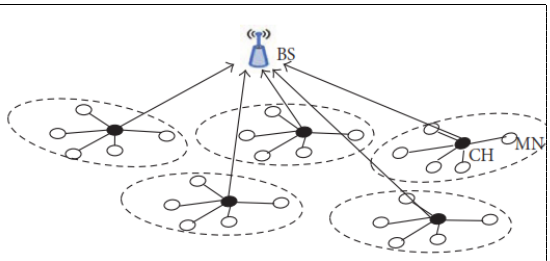

Figure 1. An example of Cluster-based WSN architecture.

Data can be transmitted over long distances via intermediate nodes as WSNs are susceptible to both internal and external outbreaks. More generally, due to their resource-restricted nature [3], they cannot deal with a hard attacker. In this situation, it is necessary to protect the system from the attackers by a secondary stage of protection, called the intrusion detection system (IDS). Efficient IDS [4] can identify the extensive attack technologies employed by the attackers.

Unfortunately, most sensor networks can be highly susceptible to attack due to WSN

features, and antagonists can probably create network traffic that can also cause a massive drop of packets or change the original message content of packets [5]. To ensure secure communication between nodes, authentication strategies are incorporated on the network. In WSNs, the safe transmission of data between nodes is significant.

In this work, IDS uses a multilayer neural perceptron system to detect the anomaly and to detect a misuse based upon the fuzzy rules. Also, the feed forward neural network is used to integrate the results of the detection and to signify the various kinds of attackers (similar to Sybil, the wormhole attack, and the hello flood attack).

Data transmission in longer distances can be performed through intermediate nodes since WSNs are vulnerable to internal and external outbreaks. Most commonly, they cannot handle a fierce attacker owing to their resource-restricted nature [3]. In this condition, a secondary stage of defence mostly called the Intrusion Detection System (IDS), is needed to protect the system from the attackers. The vast attacking techniques developed by the attackers can be detected by making use of efficient IDS [4].

Unfortunately, the majority of the sensor networks are susceptible to attacks because of WSN characteristics, and antagonists can create network traffic, which can also cause massive packet drop during the broadcasting of the packets or change the original content of the message in the packets [5]. Thus, authentication strategies are implemented in the network to ensure secure communication between the nodes. In WSNs, it is essential to carry out secure data transmission between the nodes.

In this work, IDS uses anomaly detection and misuse detection based on fuzzy rule sets along with the Multilayer Perceptron Neural Network. The Feed Forward Neural Network, along with the Backpropagation Neural Network, is utilized to integrate the detection results and indicate the different types of attackers (i.e., Sybil attack, wormhole attack, and hello flood attack).

The rest of the paper is organized as follows. Section 2 deals with the related work on IDS methods in WSN. Section 3 describes the methods proposed for defining the MRI brain tumor classification with DBN. Section 4 discusses the experimental findings. Section 5 includes the conclusion and prospective work.

II. RELATED WORK

For the securement of wireless sensor networks in routing attacks, a new intrusion detection framework is proposed in [6]. The proposed method focuses on the neighboring nodes in a distributed environment to identify intrusions.

A lightweight, energy-efficient system proposed in [7] that uses mobile agents for metrically detecting intrusions based on sensor nodes energy consumption. To predict energy consumption, a linear regression model is applied.

An IDS framework is inspired by the HIV system that can be used in the network of wireless sensors [8]. The Dendritic cell algorithm uses an improved decentralized, and personalized version enables nodes to monitor their area and work together to identify an intrusive.

A WSN intrusion-detection system based on the number of active, successful deliveries is being proposed to trust-based adaptive tracking acknowledgment (TRAACK) and the Kalman filter to anticipate node trust in [9]. A novel trust management scheme based on the theory of Dempster – Shafer evidence for malicious node detection is proposed to deal with the situation of the quantification and uncertainty of trust [10].

In [11], the trust system for physical sensor networks (WSNs) on the physical layer has been proposed for intrusion detection based on physical layouts (PL-IDS). In [12], the WSNs propose a game theory-based multi-level detection frame. A mixture of specification rules and a lightweight anomaly detection module based on the neural network is applied to the proposed framework for identifying malicious sensor nodes. Calculation and energy-intensive, which negatively impact on the overall lifetime of the WSNs, are the most current frameworks for the intruder detection of wireless sensor networks (WSNs). This work consequently proposes IDS for neural network detection of different attacks in WSN.

III. PROPOSED METHODOLOGY

The work being proposed is intended to detect hello flooding, wormhole, and Sybil attacks in the WSN with the IDS. To recognize attackers of different types, use an enhanced LEACH protocol (with fuzzy rules). For the detection of these attacks, IDS benefits from both anomaly detection and misuse detection models. A higher detection rate and a low positive rate may be achieved with the proposed IDS. In the meantime, MPNN's machine-learning strategy can discover, which included new instances practically by enduring unknown

attacks. The architecture diagram of IDS using MPNN is shown in Figure 1.

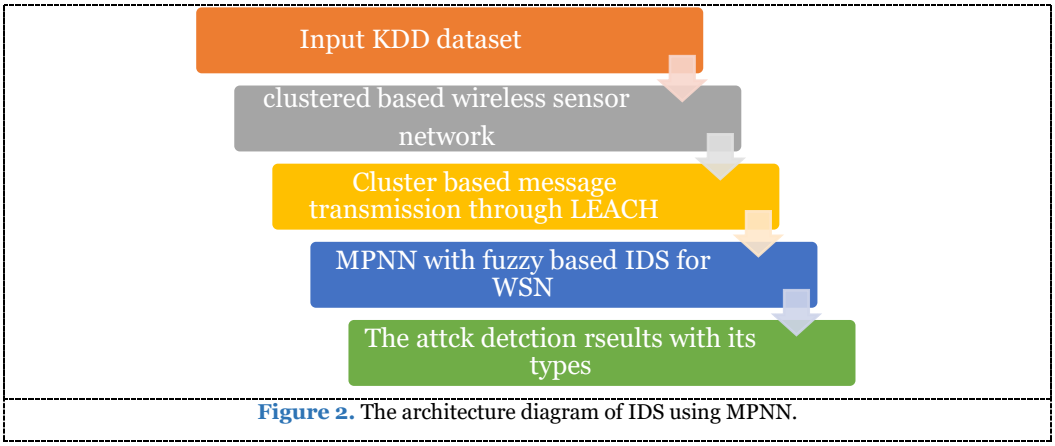

Figure 2. The architecture diagram of IDS using MPNN.

3.1. *KDD Dataset Description*

The cup KDD-99 is the most common dataset for algorithm training. It is the DARPA-98 dataset sub-set. Dataset KDD-99 is a multiform dataset. There are 4.8 million cases in this dataset. Categorical and integer characteristics of attributes are used in the dataset, and it has 42 features [12].

3.2. *Multilayer Perceptron Neural Network (MPNN)*

The model MPNN is divided into FFNN and BPNN and is used to estimate the accuracy of the detection of the three different attackers. Because of the development of advanced technology, attackers improve with each day, and thus there is a great need for the improvement of existing IDS and system capacity. The IDS is an advanced intelligent detection system to overcome such problems. The mechanism for the machine learning mechanism can identify and learn new types of attacks. However, the data packets cannot be correctly sorted by an unknown assault by the detection model. Even farther, such data packets would also be transferred to the MPNN so that the detection system can understand and introduce the new type of attack detection system. Figure 2 provides the suggested MPNN methodology with fuzzy IDS in the WSN.

This is where the MPNN is going to develop the FFNN and BPNN IDS mechanism since these neutral networks could maintain the stability of the system with a considerable number of data and can listen to various types of attackers. FFNN would, however, make headway with detection and simultaneously estimate the new types of attack. The proposed BPNN will use the MPNN supervised learning mechanism to cluster unknown attacks that incorporate an input layer, a hidden layer, and an output layer. In the detection of excesses, a scheme cannot detect accurate data packet attacks since the layer, and the input nodes of data packets are selected by using the features selected. In the beginning, the stage is the number of output nodes. Therefore, the proposed fuzzy-based MPNN would be used to create different kinds of clusters. Thus, the results of the output nodes are improved when each output node sets a new type of detection method. An artificially supervised learning mechanism will be used to estimate the corresponding points for each output result in every data package of unknown attackers. The results of the output node will later be detected to determine the value of the output node.

When this value is higher than the alert value, then the packets inserted correspond to the output node; therefore, it is cluster-referred, and MPNN must simply modify the weights. However, if the respective output node value is much less than the warning value, the data packet is not the same as the connected weight. It therefore does not match this cluster. This data packet is not the same in the same way. The next winning node results must be found to see whether the alertness test can pass or else it will introduce a unique output node that will identify a new attack.

In addition, the sample data through the experimental simulation will be examined to obtain the true vigilance value. The information about a cluster is contained in the MPNN of the misuse detection scheme to include fresh detection classes, whereas the cluster-member values acquire the defined threshold.

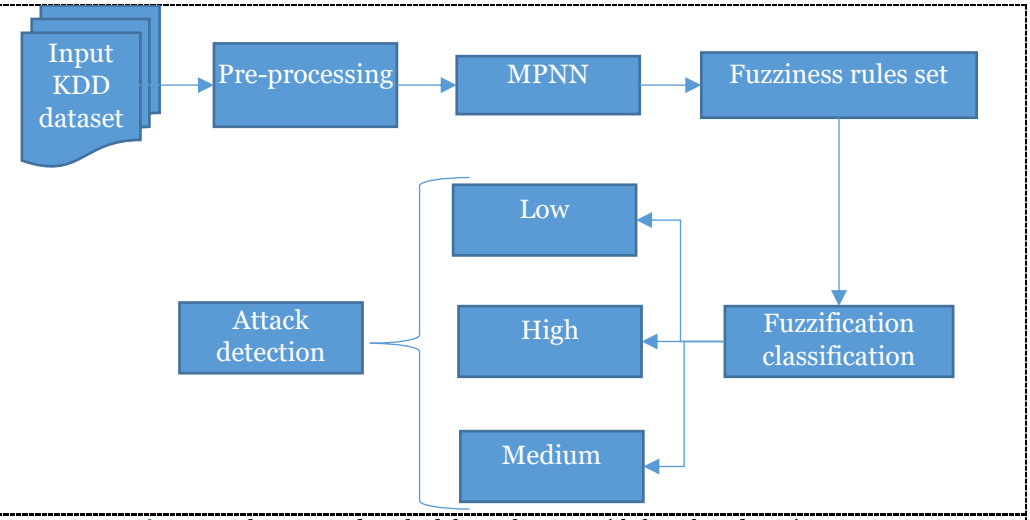

Figure 2. The proposed methodology of MPNN with fuzzy based IDS in WSN.

Fuzzy Rules Set Based Multilayer Perceptron Neural Network: The data are trained using supervised FFNN classifier by applying the n hidden nodes from the specified dataset of labelled examples, a dataset of unlabelled examples, and a test dataset. In order to get the final output as the sigmoid activation algorithm, the hidden node is implemented with BPNN supervised classification. Then, as membership vector MV is achieved on each unlabeled sample by analysing US application of MPNN supervised learning system, the membership vector of each unlabeled US sample generated during this process is further applied by using equation (1) to obtain the fuzziness F(MV):

$$F(MV) = -\frac{1}{n}\sum_{i=1}^{n}(\mu_i \log \mu_i + (1-\mu_i)\log(1-\mu_i)) * er_i \quad (1)$$

$$er_i = do_i - ao_i \quad (2)$$

where $MV = \{\mu_1, \mu_2, \ldots, \mu_n\}$ is a fuzzy set and the er_i is the error rate, do_i signifies the desired output and ao_i denotes the actual output which resulted from the MPNN. Furthermore, the Fuzziness value is categorized into three distinct groups: low, high, and medium. Samples that indicate high fuzziness and low fuzziness classes are collected, and these classes are additionally included with training darta to get a revised dataset labeled as new training data to train the FFNN and check it with BPNN.

3.3. Fuzzy with MPNN Based IDS Model for malicious node detection

Methods of anomaly detection and misuse detection use many methodologies of well-established behaviors of attack, therefore build the latest strategy to resolve or protect against these behaviors of attacks [13]. Most intrusion detection techniques promise through the training data to detect the attacks, but they fail uncertainly. The proposed analysis is based on MPNN, consisting of FFNN along with BPNN, and is implemented in this study to have the highest detection rate in the supervised approach to learning. The proposed approach showed figures outside of the traditional relationship between variables of input and output, so this fits the traditional mass. For achieve the greatest accuracy, it will minimize the error rate that exists in the code. Researchers therefore suggested fuzzy-based FFNN and BPNN to achieve the highest degree of accuracy in the detection of attacks by massive training for clustered AHIDS.

After the integration of the training data into the BPNN, can compare the actual performance results via the FFNN process. The error and rectification value of both hidden and output layers is determined by the back propagation process in the MPNN. To change the biases and weights of the networks, this length is called the epoch when all the training data have been used. The training data will be constantly discovered and periodically arranged with the aid of the epochs, the weights according to the layers, unless the output layer value is the same as the target value and then the training data is completed. Full irregular packets are then detected by the anomaly detection system, and then forwarded to the system for misuse for further verification.

Next, add the pre-processing step to cover up the irregular packets into a binary value, and then forward the binary value to the output value calculation scheme for misuse detection. Eventually, to get the best integration, the product of the detection value is provided to the fuzzy module with MPNN model. The fuzzy module is used to allow the best decision taking to classify the attackers and their different forms of attack by combining anomaly detection scheme and detection module for misuse. Through applying the rules to combine the outputs of the two detection schemes, the fuzzy rule-based method is used to support the decision-making model, and the main benefit of this

study is to obtain fast and reliable results. Tabulation gives the fuzzy dependent rules. This mechanism operates using a fuzzy logic controller; first, the input parameters (FPNN) are assigned to the Fuzzy Inference Method (FIS) in the fuzzification process. The FIS performs on the basis of the Fuzzy membership (triangular) and the Fuzzy rule that is applied to the input parameters to determine the correct fuzziness to identify the concentrations of attackers. Therefore, in FIS, these parameters are analyzed which checks the fuzzy rules and functions to produce the results to defuzzify where the output parameters are extracted as low (Sybil attack), mild (wormhole attack), and high (hello flooding attack).

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed MPNN with fuzzy is evaluated in this section, and the performance results are compared with existing TRAACK [9], PL-IDS [11] and lightweight neural network [12] schemes. The performance measurement is done in terms of precision, f-measurement, recall, accuracy and attacker detecetion rate. Here estimate the different types of attackers such as hello flooding, wormhole, and Sybil attacks.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (3):

$$Precision = \frac{TP}{FP+TP} \quad (3)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (4):

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

F-measure: this is also referred to as F 1-score, and as in equation (5) is the harmonic mean of precision and recall:

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (5)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (6):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Attacker detection rate: It is calculated using equation (7)

$$\text{Attacker detection rate} = \sum_{i=2}^{c} \frac{TP_i}{TP_i+FN_i}. \quad (7)$$

Where, c is the number of classes, true positive (TP_i) samples are properly classified as no attacker of the ith class, false positive (FP) samples are incorrectly classified as attacker, True negative (TN) samples are properly classified as attacker, and false negatives (FN_i) are incorrectly classified as attacker of the ith class.

4.1. *Precision Rate comparison*

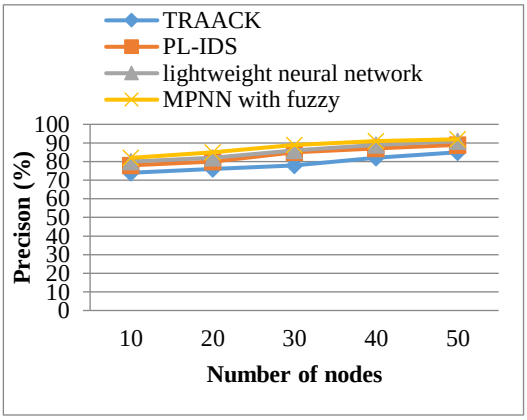

Figure 3. Representation of Precision Comparison

From the above Figure 4, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented TRAACK, PL-IDS, lightweight neural network and MPNN with fuzzy. When the number of records increases according to the precision value, from this graph, it is learned that the proposed MPNN with fuzzy offers 92% higher precision than previous methods that yield better results in the classification of attackers. The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

Number of nodes	TRAACK	PL-IDS	lightweight neural network	MPNN with fuzzy
10	74	78	80	82
20	76	80	82	85
30	78	85	86	89
40	82	87	89	91
50	85	89	91	92

4.2. *Recall comparison*

From the above Figure 5, the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as TRAACK, PL-IDS, lightweight neural network and MPNN with fuzzy. Increasing the number of images often increases the correct value for the recall. Through this graph, it is discovered that the current MPNN with fuzzy offers recall 95% higher than previous methods.

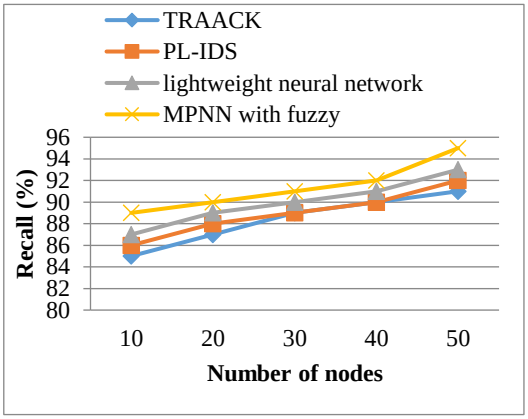

Figure 4. Representation of Recall Comparison

The explanation for this is that the MPNN with fuzzy classifies the fuzziness, which will enhance the detection and classification of attackers in WSN. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

Number of nodes	TRAACK	PL-IDS	lightweight neural network	MPNN with fuzzy
10	85	86	87	89
20	87	88	89	90
30	89	89	90	91
40	90	90	91	92
50	91	92	93	95

4.3. F-measure Rate comparison

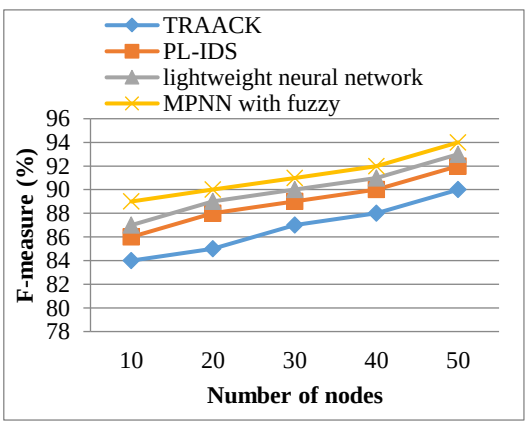

Figure 5. Representation of F-measure Comparison

From the above Figure 6, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as TRAACK, PL-IDS, lightweight neural network and MPNN with fuzzy. When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed MPNN with fuzzy offers 95% higher f-measurement than previous methods. Therefore the proposed MPNN with fuzzy algorithm is stronger than the current algorithms in terms of better performance of classifying attackers in WSN. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

Number of nodes	TRAACK	PL-IDS	lightweight neural network	MPNN with fuzzy
10	85	89	90	91
20	87	90	91	92
30	89	91	92	93
40	90	92	93	94
50	91	93	94	95

4.4. Accuracy comparison

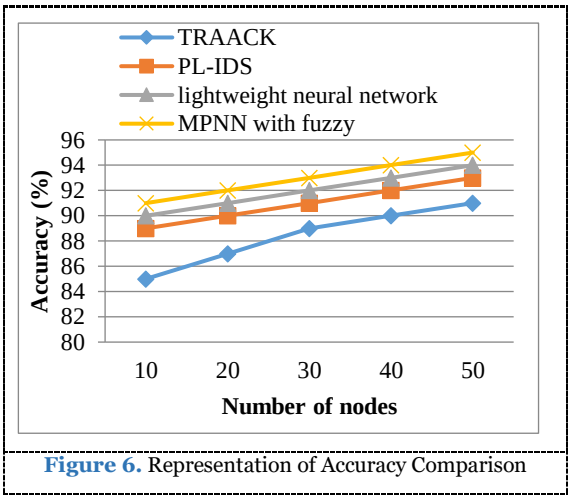

Figure 6. Representation of Accuracy Comparison

From the above Figure 7, the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as TRAACK, PL-IDS, lightweight neural network and MPNN with fuzzy.

From the graph, it is known that the proposed MPNN with fuzzy algorithm is higher than the existing algorithms with a high precision rate of 95% in terms of better attacking results. This is due to the fuzzy based extraction of the fuzziness in the MPNN algorithm, which increases the attacker classification results. The numerical results of Accuracy Comparison is shown in Table 4.

Table 4. The numerical results of Accuracy Comparison

Number of nodes	TRAACK	PL-IDS	lightweight neural network	MPNN with fuzzy
10	85	89	90	91
20	87	90	91	92
30	89	91	92	93
40	90	92	93	94
50	91	93	94	95

4.5. Attacker detection rate comparison

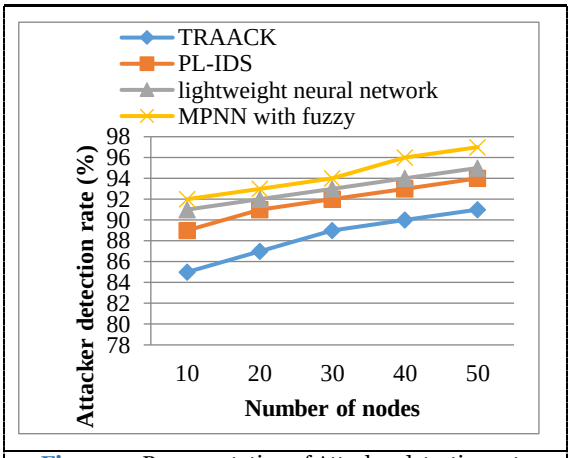

Figure 7. Representation of Attacker detection rate Comparison

From the above Figure 7, the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as TRAACK, PL-IDS, lightweight neural network and MPNN with fuzzy. From this graph, it is known that the proposed MPNN with fuzzy algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better attacker classification results. The numerical results of Accuracy Comparison is shown in Table 5.

Table 5. The numerical results of Attacker detection rate Comparison

Number of nodes	TRAACK	PL-IDS	lightweight neural network	MPNN with fuzzy
10	85	89	91	92
20	87	91	92	93
30	89	92	93	94
40	90	93	94	96
50	91	94	95	97

V. CONCLUSION AND FUTURE WORK

In this work, an IDS against hello flooding, wormholes and Sybil attacks in wireless sensor networks is suggested for wireless sensor networks that are using both anomaly detection and misuse detection to block malicious This system utilizes a Multilayer Perceptron Neural Network based on a fuzzy logic mechanism with a high detection rate and a low false positive rate. The detection mechanism is included in LEACH protocol for cluster-based topology to minimize transmission costs and energy usage, leading to an improvement in network life. The simulation results demonstrate that a proposed MPNN is able to produce high real-life positives and low false positives. This is accomplished with a fuzzy algorithm. In order to evaluate the performance of the proposed research and also encourage biological methods to be integrated into classification algorithms, further analyses are required in the future on this subject.

REFERENCES

[1]. Sarigiannidis, P., Karapistoli, E., & Economides, A. A. (2015). Detecting Sybil attacks in wireless sensor networks using UWB ranging-based information. *expert systems with applications*, *42*(21), 7560-7572.

[2]. Patel, R., Pariyani, S., & Ukani, V. (2011). Energy and throughput analysis of hierarchical routing protocol (LEACH) for wireless sensor network. *International Journal of Computer Applications*, *20*(4), 32-36.

[3]. Sarigiannidis, P., Karapistoli, E., & Economides, A. A. (2015). Detecting Sybil attacks in wireless sensor networks using UWB ranging-based information. *expert systems with applications*, *42*(21), 7560-7572.

[4]. Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, *41*(4), 1690-1700.

[5]. Shen, Y., Liu, S., & Zhang, Z. (2015). Detection of hello flood attack caused by malicious cluster heads on LEACH protocol. *International Journal of Advancements in Computing Technology*, *7*(2), 40.

[6]. Farooqi, A. H., Khan, F. A., Wang, J., & Lee, S. (2013). A novel intrusion detection framework for wireless sensor networks. *Personal and ubiquitous computing*, *17*(5), 907-919.

[7]. Riecker, M., Biedermann, S., El Bansarkhani, R., & Hollick, M. (2015). Lightweight energy consumption-based intrusion detection system for wireless sensor networks. *International Journal of Information Security*, *14*(2), 155-167.

[8]. Salmon, H. M., De Farias, C. M., Loureiro, P., Pirmez, L., Rossetto, S., Rodrigues, P. H. D. A., ... & da Costa Carmo, L. F. R. (2013). Intrusion detection system for wireless sensor networks using danger theory immune-inspired techniques. *International journal of wireless information networks*, *20*(1), 39-66.

[9]. Rajeshkumar, G., & Valluvan, K. R. (2017). An energy aware trust based intrusion detection system with adaptive acknowledgement for wireless sensor network. *Wireless Personal Communications*, *94*(4), 1993-2007.

[10]. Zhang, W., Zhu, S., Tang, J., & Xiong, N. (2018). A novel trust management scheme based on Dempster–Shafer evidence theory for malicious nodes detection in wireless sensor networks. *The Journal of Supercomputing*, *74*(4), 1779-1801.

[11]. Ghugar, U., Pradhan, J., Bhoi, S. K., Sahoo, R. R., & Panda, S. K. (2018). PL-IDS: physical layer trust based intrusion detection system for wireless sensor networks. *International Journal of Information Technology*, *10*(4), 489-494.

[12]. Ravipati, R. D., & Abualkibash, M. (2019). Intrusion Detection System Classification Using Different Machine Learning Algorithms on KDD-99 and NSL-KDD Datasets-A Review Paper. *International Journal of Computer Science & Information Technology (IJCSIT) Vol*, *11*.

[13]. Yan, K. Q., Wang, S. C., Wang, S. S., & Liu, C. W. (2010, July). Hybrid intrusion detection system for enhancing the security of a cluster-based wireless sensor network. In *2010 3rd international conference on computer science and information technology* (Vol. 1, pp. 114-118). IEEE.

Cite this article as: Narmatha C. A New Neural Network-Based Intrusion Detection System for Detecting Malicious Nodes in WSNs. J. Comput. Sci. Intell. Technol. 2020; 1(3): 01–08. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 3, pages: 09 –14, December 2020

A Proficient Adaptive K-means based Brain Tumor Segmentation and Detection Using Deep Learning Scheme with PSO

[1]Anitha T, [1]Charlyn Pushpa Latha G, & [2]Surendra Prasad M

[1]Department of Computer Science & Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Deemed to be University), Chennai, Tamil Nadu, India.

[2] Annapoorana Medical College and Hospitals. Salem - Kochi Highway, Kombadipatti, Tamil Nadu, India

***Corresponding Author: anitha.bioinformatics@gmail.com*

Received: *10.08.2020,*
Revised: *14.11.2020,*
Accepted: *14.12.2020,*
Published: *22.12.2020*

DOI:
10.53409/mnaa.jcsit20201302

Abstract: Determining the size of the tumor is a significant obstacle in brain tumour preparation and objective assessment. Magnetic Resonance Imaging (MRI) is one of the non-invasive methods that has emanated without ionizing radiation as a front-line diagnostic method for brain tumour. Several approaches have been applied in modern years to segment MRI brain tumours automatically. These methods can be divided into two groups based on conventional learning, such as support vector machine (SVM) and random forest, respectively hand-crafted features and classifier method. However, after deciding hand-crafted features, it uses manually separated features and is given to classifiers as input. These are the time consuming activity, and their output is heavily dependent upon the experience of the operator. This research proposes fully automated detection of brain tumor using Convolutional Neural Network (CNN) to avoid this problem. It also uses brain image of high grade gilomas from the BRATS 2015 database. The suggested research performs brain tumor segmentation using clustering of k-means and patient survival rates are increased with this proposed early diagnosis of brain tumour using CNN.

Keywords: Brain tumor, Computer-Aided Diagnosis, Deep Learning, Fusion Feature, Recurrent Extreme Learning Machine, Artificial Bee Colony.

I. INTRODUCTION

Brain tumors are a heterogeneous group of neoplasms of the central nervous system that form within or adjacent to the brain. The incidence of tumors in India's central nervous system (CNS) ranges from 5 to 10 per 100,000 population with a rising trend [1]. Metastatic brain tumors arise when cancer travels to the brain situated in another organ of the body. Forty per cent of all cancers have spread to the brain. Recently computer-aided techniques have been investigated for the analysis and visualization of magnetic resonance (MR) images. A lot of researchers have focused on detecting and quantifying brain abnormalities. A significant step in this process is the automatic identification of the brain in head MR images. Another essential move for computer-aided research is the assurance of the data quality. The MR images involve unintended differences in intensity due to MRI scanner imperfections.

Removing or reducing such variations will increase automated analysis accuracy. This research introduces a new, fully automated method for the detection of intracranial boundary and intensity correction in head MR images. The intracrane boundary is the boundary between the intracranial cavity and the brain. It separates the brain correctly from other features within the head.

A new approach based on CNN is presented in this research which enables automatic detection and segmentation of brain tumors. The approach is based on the geodesic distance and covariance. The covariance approach for detecting central coordinates of irregular tissues is based on. Using adaptive clustering of k-means, these coordinates are used for segmenting the brain tumor region. The ultimate goal is to retrieve the tumor attributes observed on the image in order to use them in the segmentation and classification stage. The present methods are based on MR images and have shown a better performance in biomedical image analysis.

The structuring of the rest of paper is as follows: In Section 2, discuss brain tumor detection methods. Section 3 describes the proposed

methodology. Section 4 describes the results and discussion; in Section 5, conclude the paper and define the future work.

II. RELATED WORK

A hybrid brain tumor detection algorithm using statistical features in magnetic resonance images and a Fuzzy including kernel support vector machine (FSVM) classifier is proposed in [2]. Brain tumors are not diagnosed early and are not fully treated and patients may suffer permanent brain injury or death. Place and size of tumors are critical to effective treatment.

In the proposed system [3], the self-organizing map neural network initially trains the features extracted from the discrete wavelet transform blend wavelets and subsequently, the resulting filter factors are equipped by the K-nearest neighbor (KNN). The aim of the study in [4] is to use advanced image processing techniques and probabilistic neural network (PNN) method to detect and locate tumor areas in the brain.

An unsupervised tumor segmentation clustering method is proposed at [5]. In addition, a fused feature vector is used which is a mixture of the features of Gabor wavelet (GWF), oriented gradient histograms (HOG), local binary pattern (LBP) and fractal texture analysis based segmentation (SFTA). Random forest (RF) classifier is used to identify three sub-tumor regions such as the total, enhancing, and non-enhancing tumor.

In [6] Neural Network Ensemble is used to improve system accuracy. The Jaya algorithm used for segmentation, and also extraction of the abnormal brain section, requires less time to execute compared to the Particle Swarm Optimization and Genetic Algorithm used previously.

The fuzzy logic-based hybrid kernel is built in [7] and used to train the vector support system to automatically identify four different forms of brain tumors such as Meningioma, Glioma, Astrocytoma, and Metastases.

Gray-level co-occurrence matrix (GLCM) features, which segmentation of the DWT-based brain tumor area in [8] is used to minimize complexity and improve performance. In brain MRI images the probabilistic neural network classifier was used to detect tumor location.

The proposed method in [9] classifies brain MRI slices as regular or abnormal, using Gabor filter and supporting vector machines. The problem of auto classification of brain MRI slices as normal and abnormal has been identified to resolve numerous schemes, but precision, robustness, and optimization are still an open question..

III. PROPOSED METHODOLOGY

In this method, segmentation has been combined using adaptive clustering of K-means followed by detection as a classifier tool with convolutionary neural network (CNN). This adaptive k-means method incorporates the maximum connected domain algorithm for the adaptive determination of K values of the K-means segmentation process. For a broad database, the segmented area from the adaptive k-means was categorized as the normal and abnormal tumor cells of medical brain MRI images. In this work, the use of Particle Swarm Optimization (PSO) in Convolutionary Neural Networks (CNNs), which is one of the basic methods in deep learning, was proposed. Using PSO in the training phase helps to optimize the outcomes of solution vectors on CNN to increase the accuracy of tumor identification. The result includes the fact that it makes it easier for clinical experts to make a decision about diagnosis and even scanning with this proposed process. The work's principal contribution is as follows:

- Initially, pre-processing improves image quality by removing noise and the artifacts of the input MRI image, this study employs a median filtering technique to pre-process the image.
- Image segmentation is performed using adaptive k-means clustering to partition MRI image into mutually exclusive and exhausted regions, so that each area of interest is spatially contiguous and the pixels within the area are homogeneous to the predetermined rules.
- Finally, automatic brain tumor detection is achieved using the classification of Particle Swarm Optimization in the Convolutionary Neural Networks (CNN) as shown in Figure 1.

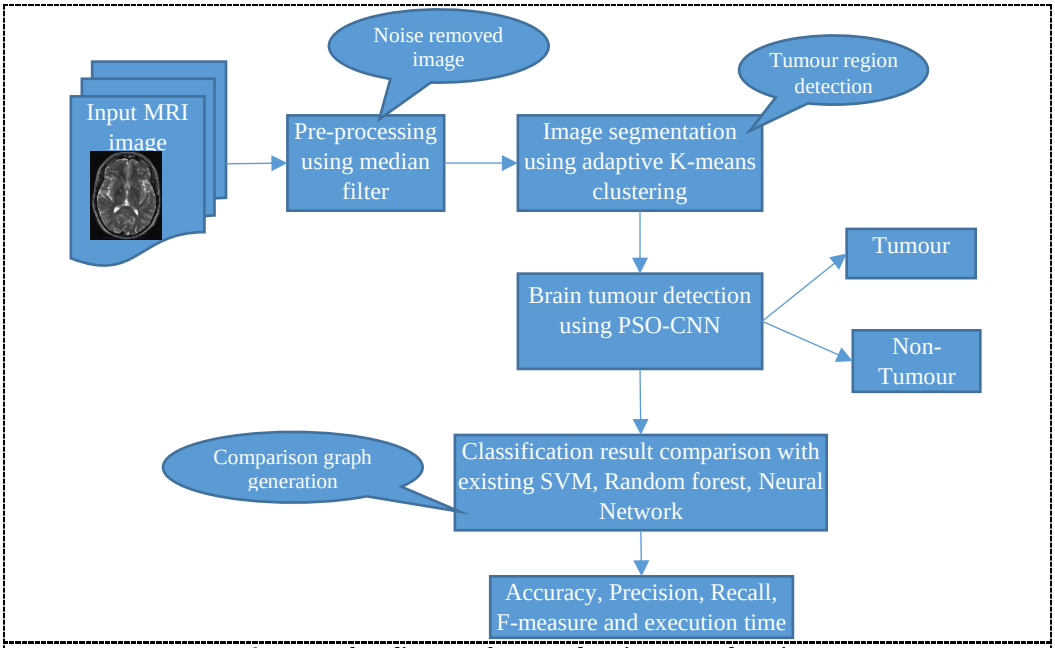

Figure 1. Flow diagram of Proposed Brain tumour detection.

3.1. Input MRI dataset

The MRI data sets were collected from the public database (http:/adni.loni.usc.edu/) of the Harvard Medical School Architecture and Alzheimer's disease Neuroimaging Initiative (ADNI). A total of 100 MR images were used for preparation, and 25 MR images were used for research. The input MRI images will experience the gray image conversion process, tumor position detection and tumor segmentation may experience correlation computation. For relegation of benign and malignant tumors using GLCM (Grey Level Co-occurrence Matrix) technique a minimum of ten features are collected.

3.2. Image Pre-processing Using Median Filter

The image processing is really complicated. Before any picture is processed, removing unnecessary objects that it may contain is quite important. The picture can be processed successfully after extracting unwanted artefacts. Image Pre-Processing is the key step to image processing. It includes processes such as grayscale image conversion, noise reduction, and image restoration. The most popular method of pre-processing is conversion to gray scale image. After the image is converted to grayscale, then using Median Filter methods to eliminate excess noise. This most famous technique used during the reduction of noise. It is a filtering method called "non-linear." It has been used to suppress the grayscale picture created by salt and pepper noise. Median filter is based on average pixel size, and therefore retains the edges and boundaries [10].

3.3. Image segmentation using adaptive K-means clustering

In this method, brain image segmentation using K-means clustering technique to segment the tumor from the image as a contour to Variational Level sets, which in turn performs tumor boundary removal. K-Means is one of the unsupervised cluster learning methods. Clustering the image groupes the pixels by the same characteristics [11].

1. Initialize k as the cluster value no.
2. Always choose k-cluster centers at random
3. Determine cluster mean or centre
4. Compute the difference between pixels and centers of each cluster
5. If the distance to the center is near then transfer to that cluster.
6. Move to next cluster otherwise.
7. Re-evaluate the centre.
8. Repeat the cycle before the core comes in.

3.4. Feature Extraction using GLCM

Features of the image were extracted using the Gray Level Co-occurrence Matrix (GLCM). In this feature extraction technique, initially GLCM of the image was calculated then the features based on the GLCM were extracted. The extracted features include contrast, correlation, energy, homogeneity, mean, standard deviation, entropy, skewness, kurtosis, inverse difference moment.

3.5. *Brain tumour detection using PSO-CNN*

In this study, we perform a combination of biologically inspired Particle Swarm Optimization and Convolutional Neural Network as a classifier tool to improve diagnostic accuracy. The cause of this study is to extract information from the segmented tumor region and classify healthy and infected tumor tissues for a large database of medical images. Our results lead to conclude that the proposed method is suitable to integrate clinical decision support systems for primary screening and diagnosis by the radiologists or clinical experts.

Convolutional Neural Networks (CNN) are one of the most used neural networks in the present time. Its applications are extremely varied. Most recently they have been proving helpful with deep learning, as well. Since it is growing in more convoluted domains, its training complexity is also increasing. To tackle this problem, many hybrid algorithms have been implemented. In this paper, Particle Swarm Optimization (PSO) is used to reduce the overall complexity of the algorithm. The hybrid of PSO used with CNN decreases the required number of epochs for training and the dependency on GPU system. The algorithm so designed is capable of achieving 3-4% increase in accuracy with lesser number of epochs. The advantage of which is decreased hardware requirements for training of CNNs. The hybrid training algorithm is also capable of overcoming the local minima problem of the regular backpropagation training methodology.
Generally, the process of the proposed method consists of several steps as shown below:

1) The first step is initializing the learning rate of the CNNs with the value is 1 based on the experiment. Batch size of CNNs is 50, the number of CNNs epoch in the range of 1 to 4, PSO iteration is 10. The convergence status of PSO is used to check the convergences of PSO, if the error value has not changed for three iterations, then the PSO is considered as convergent;
2) After setting up the experiment, the next step is run CNNs training process, where the tumour is detected from the feature extraction process. The result of CNNs is vector output that will be optimizing using PSO algorithm. PSO optimization in this study serves to make the value of loss function on CNN becomes minimal;
3) The output vector will be update if the solution of swarm has less error compare with old vector output;
4) The PSO will run as long as the iteration number of PSO and the convergence solution have not fulfilled;
5) After the CNN Training, the model will be tested with testing data that consist of MRI brain image data;
6) The result of CNN test is accuracy of CNN, it represent how precise of the CNN model can predict the actual value of testing dataset.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed PSO-CNN is evaluated in this section, and the performance results are compared with existing FSVM [1], KNN [2] and PNN [3] image compression schemes. The facts given below show that the device proposed has achieved better performance in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (6):

$$Precision = \frac{TP}{FP+TP} \quad (6)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (7):

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

F-measure: this is also referred to as F 1-score, and as in equation (8) is the harmonic mean of precision and recall:

(8):

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (8)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (9):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

Where true positive (TP) samples are properly classified as natural, false positive (FP) samples are incorrectly classified as irregular, True negative (TN) samples are properly classified as irregular, and false negatives (FN) are incorrectly classified as natural.

4.1. *Precision Rate comparison*

From the above Figure 4, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as FSVM, KNN, PNN, and PSO-CNN. When the number of records increases according to the precision value.

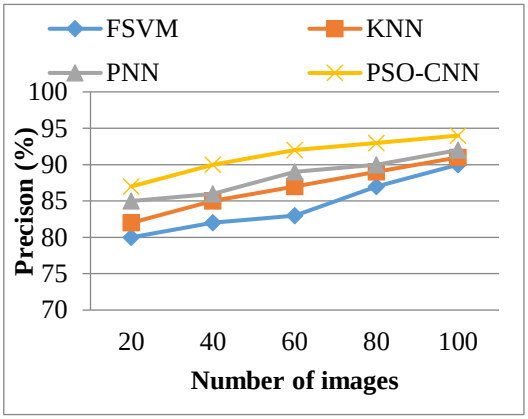

Figure 4. Representation of Precision Comparison

From this graph, it is learned that the proposed PSO-CNN offers 94% higher precision than previous methods that yield better results in the classification of Brain tumour mass due to prior segmentation of the Brain tumour mass using k-means technique. The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

No. of images	FSVM	KNN	PNN	PSO-CNN
20	80	82	85	87
40	82	85	86	90
60	83	87	89	92
80	87	89	90	93
100	90	91	92	94

4.2. Recall comparison

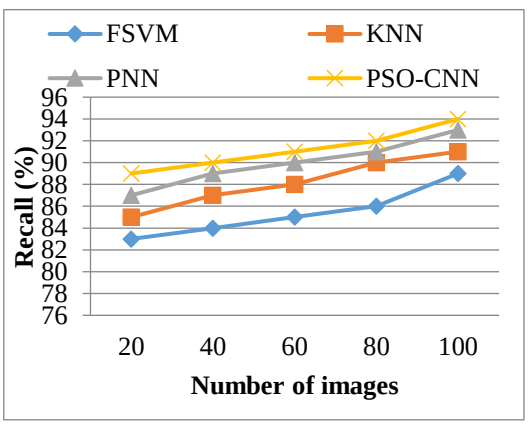

Figure 5. Representation of Recall Comparison

From the above Figure 5 the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as FSVM, KNN, PNN, and PSO-CNN. Increasing the number of photographs often increases the correct value for the recall. Through this graph, it is discovered that the current PSO-CNN offers recall 94% higher than previous methods. The explanation for this is that the PSO-CNN extracts the features directly which will enhance the Brain tumour classification tests. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

No. of images	FSVM	KNN	PNN	PSO-CNN
20	83	85	87	89
40	84	87	89	90
60	85	88	90	91
80	86	90	91	92
100	89	91	93	94

4.3. F-measure Rate comparison

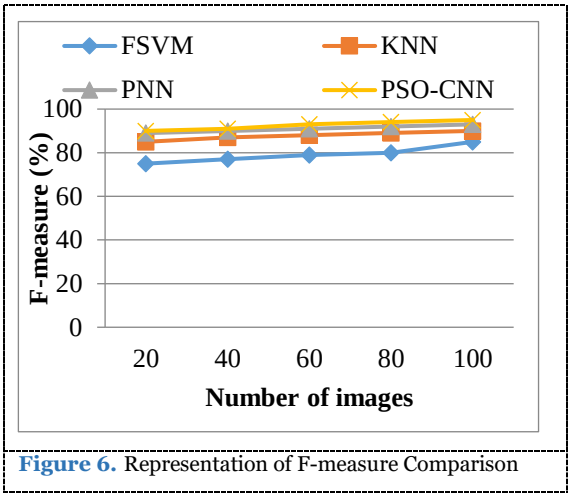

Figure 6. Representation of F-measure Comparison

From the above Figure 6, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as FSVM, KNN, PNN, and PSO-CNN. When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed PSO-CNN offers 95% higher f-measurement than previous methods.

Therefore the proposed PSO-CNN algorithm is stronger than the current algorithms in terms of better performance of classifying Brain tumour. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

No.of images	FSVM	KNN	PNN	PSO-CNN
20	75	85	89	90
40	77	87	90	91
60	79	88	91	93
80	80	89	92	94
100	85	90	93	95

4.4. *Accuracy comparison*

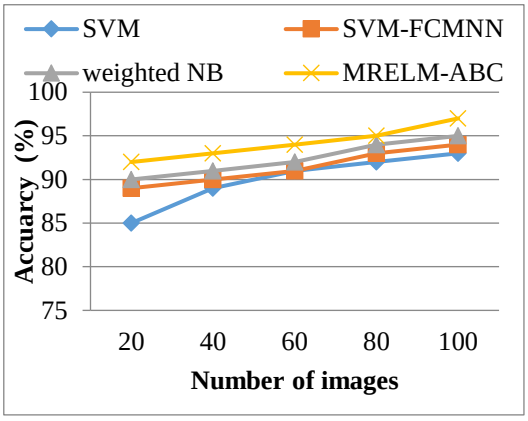

Figure 7. Representation of Accuracy Comparison

From the above Figure 7 the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as FSVM, KNN, PNN, and PSO-CNN. From this graph it is known that the proposed PSO-CNN algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better template matching results. This is due to the automatic extraction of the features using CNN in the PSO-CNN classification algorithm, which increases the classification precision resulting in Brain tumour. The numerical results of Accuracy Comparison is shown in Table 4.

Table 4. The numerical results of Accuracy Comparison

No.of images	FSVM	KNN	PNN	PSO-CNN
20	85	89	90	92
40	89	90	91	93
60	91	91	92	94
80	92	93	94	95
100	93	94	95	97

V. CONCLUSION AND FUTURE WORK

In this work, a method to extract brain tumor from 2D Magnetic Resonance brain Images (MRI) by adaptive k-means clustering algorithm is proposed which was followed by traditional classifiers and convolutional neural network. In this work, we propose a novel algorithm to search for deep convolutional neural networks (CNNs) architectures based on particle swarm optimization (PSO-CNN). A novel directly encoding strategy is also proposed in which a CNN architecture is divided into two blocks: one block contains only convolutional and pooling layers, while the other contains only fully connected layers. This encoding strategy allows for variable length CNN architectures to be compared and combined using an almost standard PSO algorithm. Our results show that PSO-CNN can quickly find an optimized CNN architecture for any given dataset. With only 30 particles and 20 iterations, the algorithm finds models capable of achieving test errors comparable to those designs exploiting more complex and complicated architectures. From the experimental results, we can conclude that psoCNN would be able to find even better architectures if more computational power is available for finding MRI brain tumor. For future works, we will expand the proposed PSO-CNN to find to include ResNets and DenseNets architectures as well, which will further increase the classification accuracy of the possible solutions.

REFERENCES

[1]. Deshpande, R. P., Babu, D., Panigrahi, M., Sekhar, Y. C., & Babu, P. P. (2016). Brain tumors incidences and a retrospective clinical analysis from a tertiary hospital in India. *Journal of neuro-oncology*, *129*(2), 383-387.

[2]. Jayachandran, A., & Sundararaj, G. K. (2015). Abnormality segmentation and classification of multi-class brain tumor in MR images using fuzzy logic-based hybrid kernel SVM. *International Journal of Fuzzy Systems*, *17*(3), 434-443.

[3]. Anitha, V., & Murugavalli, S. J. I. C. V. (2016). Brain tumour classification using two-tier classifier with adaptive segmentation technique. *IET computer vision*, *10*(1), 9-17.

[4]. Ural, B. (2018). A computer-based brain tumor detection approach with advanced image processing and probabilistic neural network methods. *Journal of Medical and Biological Engineering*, *38*(6), 867-879.

[5]. Amin, J., Sharif, M., Raza, M., & Yasmin, M. (2018). Detection of brain tumor based on features fusion and machine learning. *Journal of Ambient Intelligence and Humanized Computing*, 1-17.

[6]. Kaur, K., Walia, G. K., & Kaur, J. (2018). Neural Network Ensemble and Jaya Algorithm Based Diagnosis of Brain Tumor Using MRI Images. *Journal of the Institution of engineers (India): Series B*, *99*(5), 509-517.

[7]. Jayachandran, A., & Sundararaj, G. K. (2015). Abnormality segmentation and classification of multi-class brain tumor in MR images using fuzzy logic-based hybrid kernel SVM. *International Journal of Fuzzy Systems*, *17*(3), 434-443.

[8]. Shree, N. V., & Kumar, T. N. R. (2018). Identification and classification of brain tumor MRI images with feature extraction using DWT and probabilistic neural network. *Brain informatics*, *5*(1), 23-30.

[9]. Gilanie, G., Bajwa, U. I., Waraich, M. M., Habib, Z., Ullah, H., & Nasir, M. (2018). Classification of normal and abnormal brain MRI slices using Gabor texture and support vector machines. *Signal, Image and Video Processing*, *12*(3), 479-487.

[10]. Roy, S., & Bandyopadhyay, S. K. (2012). Detection and Quantification of Brain Tumor from MRI of Brain and it's Symmetric Analysis. *International Journal of Information and Communication Technology Research*, *2*(6).

[11]. Moftah, H. M., Azar, A. T., Al-Shammari, E. T., Ghali, N. I., Hassanien, A. E., & Shoman, M. (2014). Adaptive k-means clustering algorithm for MR breast image segmentation. *Neural Computing and Applications*, *24*(7-8), 1917-1928.

Cite this article as: Anitha et al. A Proficient Adaptive K-means based Brain Tumor Segmentation and Detection Using Deep Learning Scheme with PSO. J. Comput. Sci. Intell. Technol. 2020; 1(3): 09–14. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 3, pages: 15–21, December 2020

A New Modified Recurrent Extreme Learning with PSO Machine Based on Feature Fusion with CNN Deep Features for Breast Cancer Detection

[1]Surendra Prasad M, [2]Manimurugan S

[1] Annapoorana Medical College and Hospitals. Salem - Kochi Highway, Kombadipatti, Tamil Nadu, India

[2]Faculty of Computers and Information Technology, University of Tabuk, KSA.

***Corresponding Author: semanimurugan@gmail.com*

Received: *20.08.2020,*
Revised: *14.11.2020,*
Accepted: *16.12.2020,*
Published: *22.12.2020*

DOI:
10.53409/mnaa.jcsit20201303

Abstract: Breast cancer is a prevalent cause of death, and is the only form of cancer that is common among women worldwide and mammograms-based computer-aided diagnosis (CAD) program that allows early detection, diagnosis and treatment of breast cancer. But the performance of the current CAD systems is still unsatisfactory. Early recognition of lumps will reduce overall breast cancer mortality. This study investigates a method of breast CAD, focused on feature fusion with deep features of the Convolutional Neural Network (CNN). First, present a scheme of mass detection based on CNN deep features and modified clustering of the Extreme Learning Machine (MRELM). It forecasts load through Recurrent Extreme Learning Machine (RELM) and utilizes Artificial Bee Colony (ABC) to optimize weights and biases. Second, a collection of features is constructed that relays deep features, morphological features, texture features, and density features. Third, MRELM classifier is developed to distinguish benign and malignant breast masses using the fused feature set. Extensive studies show the precision and efficacy of the proposed method of mass diagnosis and classification of breast cancer.

Keywords: Breast cancer, Computer-Aided Diagnosis, Deep Learning, Fusion Feature, Recurrent Extreme Learning Machine, Artificial Bee Colony.

I. INTRODUCTION

The malignant tumor activating in the breast cells is breast cancer (BC). A tumor is likely to spread to other body areas [1]. BC is a universal disease that typically hammers women's lives in the 25-50 age category. The potential increase in the number of BC cases in India is alarming. The BC survival rate in the last five years is approximately 90% in the United States, compared to approximately 60% in India [2]. For India in 2020, BC's projections suggest that the statistic is as high as 2 million [3].
Doctors have identified hormones, lifestyle and environmental factors that can increase the chance of developing BC of a person. More than 5%–6% of patients in BC have been associated with gene mutations through the ages of the family. The other factors that cause BC are obesity, an increasing age, postmenosal hormonal imbalances.

As such, there is no mechanism for preventing BC, however, early detection can improve the results considerably. In addition, the cost of treatment may also be significantly reduced. However, cancer symptoms can often be unusual, so it is difficult to detect them early. In order to detect any early irregularities before it develops, mammograms and self-breasts are essential [4].
The demand for machine learning nowadays increases until it is a service. Machine learning unfortunately continues to be a field of high barriers and often requires skills. A number of skills and expertise are required for an efficient machine learning model including the phases of preprocessing, selection of features and classification processes.
In addition, advances in CNNs can help radiologists as well as eventually diagnosis systems in the near future to read mammograms independently. Propose a method using CNN and MRELM for the extraction and clustering of features, respectively. First, a mammogram will be divided into various sub-regions. Then, CNN is used to extract functions according to individual sub-regions, followed by the use of MRELM with ABC in the cluster of sub-regions where the breast tumor region eventually is located. During the diagnostic period, each

mammogram used as MRELM input for classification is based on its fusion-deep features. The result directly determines if the patient has a benign or malevolent tumor of the breast. Finally, the experimental results show that the best performance is achieved by our proposed methods, the sub-regional ELM clustering and the MRELM clustering with the deep fusion features.

The structuring of the rest of paper is as follows: In Section 2, discuss breast cancer detection methods. Section 3 describes the proposed methodology. Section 4 describes the results and discussion; in Section 5, conclude the paper and define the future work.

II. RELATED WORK

In this study, the application of the K-Particle Swarm Optimization (KPSO) variation is proposed for a new hybrid for breast cancer detection. [5] KPSO is used to initialize and then update the centers and variances of the radial functional neural network with back propagation.

Texture features were removed from the co-occurrence matrix and matrix for runtime lengths and features for automatic classification of normal and maligneous breast conditions were added to the Support Vector Machine (SVM) classifier in [6].

Compared to [7] for the automatic classifying of images of breast cancer histology in two machine-learning approaches, benign and malignant subclasses. The first approach is to extract a set of man-made characteristics encoded by two coding models (wordsack and locality-restricted linear codification) and trained in the use of vector support machines.

A CAD system is developed to characterize the breast nodules as either benign, or malignant on an ultrasound image. The system aims at a high performance classifier. In [8] is developed a Fuzzy Cerebellar Model CAD Network (FCMNN).

The traditional way to diagnose the disease is based on the experience of human beings to identify certain patterns from the database. It is likely to be mistaken, to take time and to work hard. Therefore an automatic breast cancer diagnostic technique was proposed for the parameter optimization of the Artificial Neural Net Network (ANN) using a genetic algorithm (GA).

In [10], a new classification of NB (weighted NB) was proposed, with its application on detection of breast cancer submitted. Several tests on the weighted NB on the breast cancer database were conducted to evaluate its performance.

III. PROPOSED METHODOLOGY

In this research, consider five steps in the detection of breast-cancer: pre-processing of the breast image, mass detection, extraction of features, generation of data and classification training. Delusions and increased contrast processes on the original mammograms have been used in breast image pre-processing to increase the difference between the masses and the surrounding tissue. The ROI is then located for mass detection. Then the ROI extracts features such as deep characteristics, morphological characteristics, textures and density characteristics. Each image from the breast image dataset with the extracted features and corresponding labels was trained by classifiers during the training process. Thus, the diagnosed mammograms can be identified by the well-trained classifications. Fig. 1 presents the flow diagram of the entire diagnosis process.

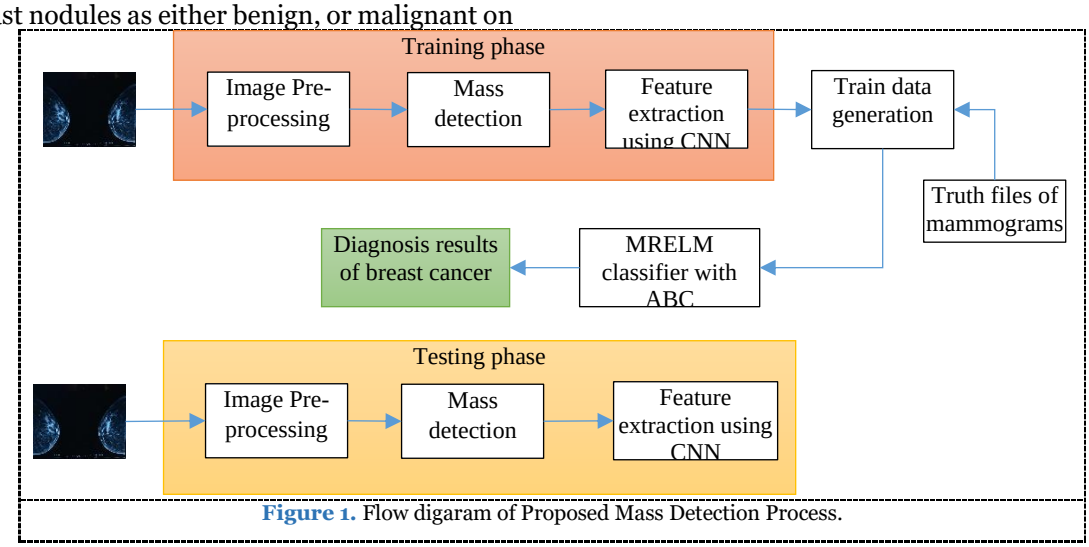

Figure 1. Flow digaram of Proposed Mass Detection Process.

3.1. Image Pre-processing

In order to avoid the effect of noise on the subsequent auxiliary diagnosis, the adaptive mean filter algorithm [11] is chosen to eliminate noise from the original mammogram. The main idea of this is to use a window with a fixed size which slides in the line direction to ascertain whether the noise is available in the window, calculating the mean, variance and spatial correlation values for each sliding window. If detect noise, start replacing the mean value of the pixel of the chosen window. In this work, an algorithm for improving the contrast between the suspected masses and the surrounding tissues has been employed. The key idea is to uniformly distribute the histogram of the original image. The gray scale of the picture is expanded following that process, while also improving the contrast and clarifying the details of the image.

3.2. Mass Detection and extract ROI

Mass detection is intended to remove the mass area from normal tissues. The more precise the mass division is, the more precise the features extracted. A mass detection process based on deep features of sub-domain CNN and ELM is proposed in this paper. After pre-processing, the first step is to take ROI from the images. The ROI is then divided by sliding window into several unsurfacing sub-regions. Determine if all sub-regions have been crossed successfully. If it is yes, the deep characteristics of the sub-regions can be extracted, otherwise the deep characteristics of the sub-regions will be clustered and the mass detection process will be completed.

There are many 0 gray-value areas in mammography, which do not have an effect on breast-CAD. The mammographic area must be separated from the whole mammogram by ROI, in order to improve mammographic processing efficiency or ensure the precise monitoring diagnosis. A mass-recovery adaptive algorithm was used in this work to extract the region of breast mass. Explicitly, in a mammogram, all rows are scanned consecutively to discover the first nonzero pixel by means of abscissa signified as $x_{firstabs}$ and the last nonzero pixel with abscissa signified as $x_{lastabs}$), then all columns are formerly scanned sequentially to find the first nonzero pixel through ordinate denoted as $y_{firstord}$ and the last nonzero pixel (with ordinate denoted as $y_{lastord}$).

In this part, a method to distribute the ROI into several non-overlapping sub-regions is planned. The searching area to manage the masses from the ROI is unchanging in a rectangular area $[x_{firstabs}, x_{lastabs}, y_{firstord}, y_{lastord}]$, wherever the length of the searching rectangular is $L = x_{lastabs} - x_{firstabs}$ and width is $W = y_{lastord} - y_{firstord}$. The rectangular penetrating area is segmented by means of a sliding window with length w and width h ($L \geq l, W \geq w$). In the rectangular searching area ($L \times W$), the sliding window ($l \times w$) is enthused with a certain step size, negotiating the searching area deprived of crossing the ROI boundary. Consequently, the ROI is alienated into several equal size ($l \times w$), non-overlapping sub-regions and such sub-regions will help as the foundation for subsequent feature extraction. In this work, the size of the sliding window is fixed as 48×48 and the searching step size is equal to 48. In conclusion, the ROI has been alienated into N non-overlapping sub-region such as $(s_1, s_2, \ldots, s_N)$.

3.3. Extract Deep Features Using CNN

This work uses CNN to extract deep characteristics from the sub-regions of ROI. The CNN input is a sub-regional image of 48 bis48 dimensions which is taken from previous steps. The first coalescing layer filters 48×48×3 images input to 12 kernels size 9×9 and receives a 40×40×12 output.

$$Conv^{k}_{(i,j)} = \sum_{u,v} W^{k,l}(u,v) \cdot input^{j}(i-u, j-v) + b^{k,l} \quad (1)$$

where $W^{k,l}$ signifies the weight of the k th kernel and $b^{k,l}$ signifies the bias of k th layer. The activation value is controlled in the range [-1,1] using tanh as the activation function.

$$Output^{k}_{(i,j)} = \tanh\left(Conv^{k}_{(i,j)}\right) \quad (2)$$

The output of the first convolution layer is related with a max-pooling layer. At that point the second and third convolution/max-pooling layer are linked to one another until take the output with size 2 × 2 ×6. The fully-connected layer has 2 × 2 × 6 = 24 neurons which are the features for the next clustering analysis.

3.4. Clustering Deep Features Using MRELM

The MRELM algorithm in this section is used to cluster deep characteristics extracted from the previous CNN architecture. There are two cluster numbers and sub-regional features are divided into two categories: suspect mass areas and unsuspecting mass areas. The effect of supervised learning on the model cannot satisfy the demand when the volume of data is small. Semi-controlled learning is therefore used to improve the effect, but can also accomplish certain tasks of clustering. The ERELM algorithm is a semi-supervised learning algorithm, so that the internal relationships of structure between the unlabeled data set can be identified.

The input is the deeper matrix X of the algorithm and the output is the cluster results of the feature. In particular, Laplacian operator LO was first built from the X training set, followed by the random generation of a cloaked layer neuron output matrix.

In particular, Laplacian L operator is first constructed from the training set X, followed by a randomly generated hidden layer neuron output matrix. If the number of hidden neuron is less than the number of input neuron, we use eq.(3) in the calculation of output weight

$$\min_{\beta \in R^{nh \times n_0}} \|w\|^2 + \lambda Tr(\beta^T H^T LOW w) \qquad (3)$$

where w represents the weights between hidden layer and output layer. Then, use the equation $(I_0 + lH^T LH)v = \gamma H^T Hv$ to compute output weights. After that, compute the embedding matrix and use k-means algorithm clustering N points into K categories.

3.5. *Modified Recurrent Extreme Learning Machine with ABC*

A new medical diagnostic framework MRELM-ABC, consisting of two main phases, has been suggested for this study. The first stage is to search for the best feature combo in the medical data by using ABC to filter out redundant and insignificant information. Three steps are the basis of the proposed procedure:

Step1: Optimal network parameter, such as network approximation feature, will be finalized, as will neuron and context neurons no. For the first time, ABC learning algorithms have been used in combination with the RELM to optimize weights and biases to enhance forecast accuracy.

Step 2: ERELM accuracy of forecasts is calculated using the fitness feature for ABC, RMSE for MAE and MSE measurement as shown below is calculated as proposed technology.

$$MSE = \frac{\sum_{t=1}^{N}[a(t)-p(t)]^2}{N} \qquad (4)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}[a(t)-p(t)]^2}{N}} \qquad (5)$$

$$MAE = \sum_{t=1}^{N}[a(t)-p(t)]^2 \qquad (6)$$

Where t is current iteration, N indicates number of iamgess, a is actual value and p is predicted value. Weights are optimized in the work using ABC technology when search algorithms with network structure and learning rate are given input data. ABC searches for the best weight and bias value.

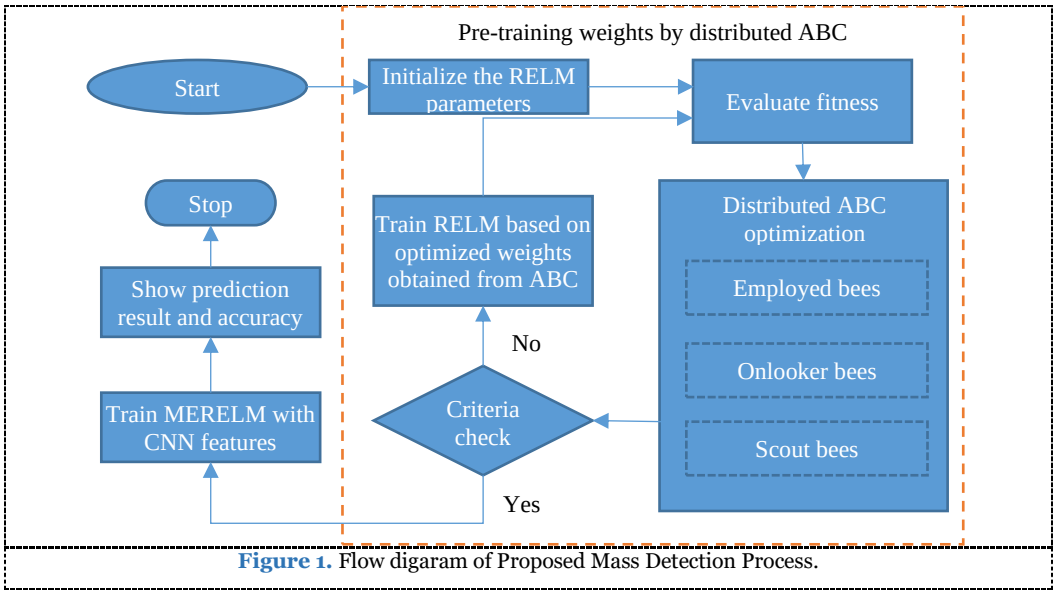

Figure 1. Flow digaram of Proposed Mass Detection Process.

The following algorithmic steps explain the mechanism for breast cancer predictions in detail.

1. Input N image sample, objective function original image dataset
2. Output Predicted Segmented part of the desired value
3. Start
4. Assign the w_i and biases b received by input weights after ABC optimisation.
5. Compute the hidden-layer output matrix H, where $H = h\ (i = 1, \dots, N)$ and $j = (1, \dots, K)$ and $h_{ij} = g(w_j \cdot x_i + b_j)$
6. Calculate the output weight matrix as $\beta = H^+T$, where H^+ shows Moore-Penrose generalized inverse of H matrix
7. Updated weights are provided for input and hidden layer as context neurons.
8. End

The MRELM algorithm is a proposed single hidden layer feeder network which has good overall performance, quick learning speed and manual

parameter setup insensitive. In this piece, MRELM was used to obtain benign and malignant diagnostic results for breast cancer.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed MRELM-ABC is evaluated in this section, and the performance results are compared with existing SVM [6], SVM-FCMNN [8] and weighted NB [10] image compression schemes. The Lung Image Database Consortium image database (LIDC-IDRI) in real time consists of diagnostic and lung cancer screening thoracic computed tomography (CT) scans of annotated marked-up lesions. Seven academic centers and eight medical imaging companies partnered to create this collection of data that contains 1018 cases. That subject contains images from a clinical thoracic CT scan and an accompanying XML file that documents the findings of four experienced thoracic radiologists conducting a two-phase image annotation procedure. The figures given below show that the device proposed has achieved better performance in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (6):

$$Precision = \frac{TP}{FP+TP} \quad (6)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (7):

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

F-measure: this is also referred to as F 1-score, and as in equation (8) is the harmonic mean of precision and recall:

(8):

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (8)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (9):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

Where true positive (TP) samples are properly classified as natural, false positive (FP) samples are incorrectly classified as irregular, True negative (TN) samples are properly classified as irregular, and false negatives (FN) are incorrectly classified as natural.

4.1. Precision Rate comparison

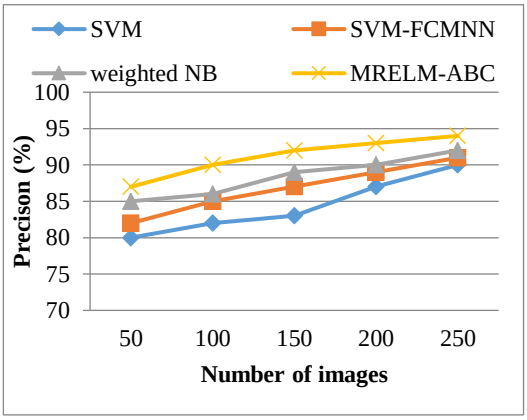

Figure 4. Representation of Precision Comparison

From the above Figure 4, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. When the number of records increases according to the precision value. From this graph, it is learned that the proposed MRELM-ABC offers 94% higher precision than previous methods that yield better results in the classification of Breast cancer mass due to prior segmentation of the Breast cancer mass using k-means technique. The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	80	82	85	87
100	82	85	86	90
150	83	87	89	92
200	87	89	90	93
250	90	91	92	94

4.2. Recall comparison

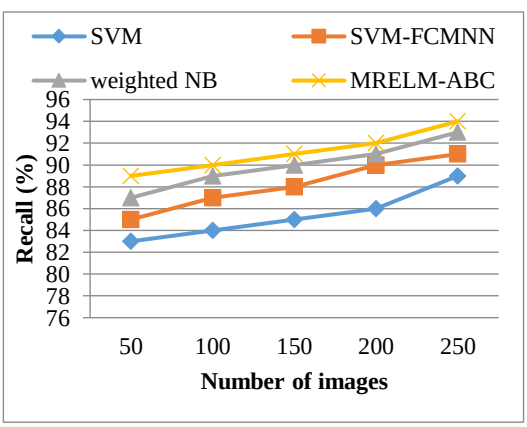

Figure 5. Representation of Recall Comparison

From the above Figure 5 the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. Increasing the number of photographs often increases the correct value for the recall. Through this graph, it is discovered that the current MRELM-ABC offers recall 94% higher than previous methods. The explanation for this is that the MRELM-ABC extracts the features directly which will enhance the Breast cancer classification tests. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	83	85	87	89
100	84	87	89	90
150	85	88	90	91
200	86	90	91	92
250	89	91	93	94

4.3. F-measure Rate comparison

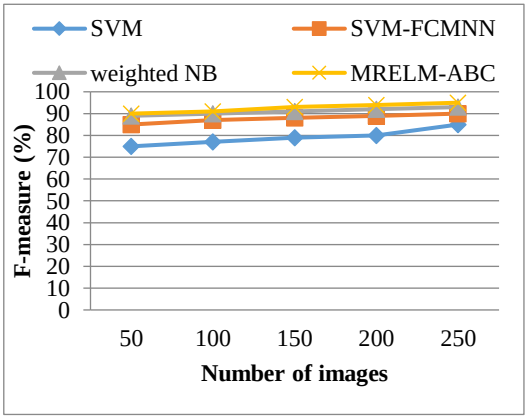

Figure 6. Representation of F-measure Comparison

From the above Figure 6, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed MRELM-ABC offers 95% higher f-measurement than previous methods.

Therefore the proposed MRELM-ABC algorithm is stronger than the current algorithms in terms of better performance of classifying Breast cancer. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	75	85	89	90
100	77	87	90	91
150	79	88	91	93
200	80	89	92	94
250	85	90	93	95

4.4. Accuracy comparison

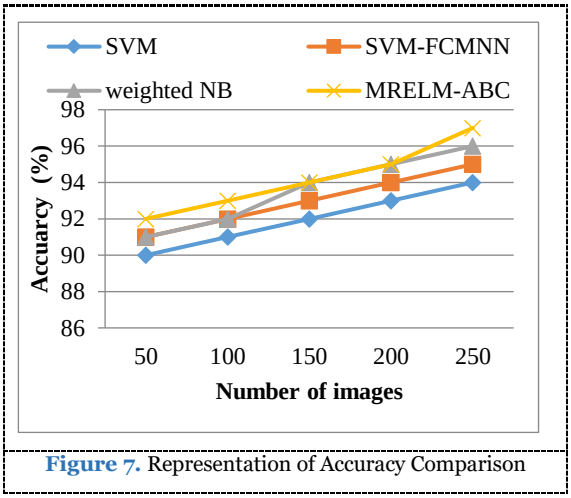

Figure 7. Representation of Accuracy Comparison

From the above Figure 7 the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as SVM, SVM-FCMNN, weighted NB, and MRELM-ABC. From this graph it is known that the proposed MRELM-ABC algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better template matching results. This is due to the automatic extraction of the features using CNN in the MRELM-ABC classification algorithm, which increases the classification precision resulting in Breast cancer. The numerical results of Accuracy Comparison is shown in Table 4.

Table 4. The numerical results of Accuracy Comparison

No.of images	SVM	SVM-FCMNN	weighted NB	MRELM-ABC
50	90	91	91	92
100	91	92	92	93
150	92	93	94	94
200	93	94	95	95
250	94	95	96	97

V. CONCLUSION AND FUTURE WORK

This work suggested a CAD breast based on deep fusion characteristics. The principal idea is the application to two phases of mass detection and mass diagnosis deep CNN extracted features. A method based on deep sub-domain CNN

characteristics and clustering is developed at the stage of mass detection. A MRELM-ABC classifier is used to classify the benign and malignant breast masses with a fused characteristic set which combines deep characteristics, morphological characteristics, texture characteristics, and density characteristics. In breast CAD, diagnostic accuracy is determined by the choice of characteristics. The classifier is used to classify the benign and malignant of the breast mass after the characteristics have been extracted. MRELM is selected as the classifier in this work, which has a better effect on multi-dimensional classification. In future work the proposed methodology will be applied to more practical issues and will plan the parallel implementation of the soft computing method using high performance tools.

REFERENCES

[1]. Hotko, Y. S. (2013). Male breast cancer: clinical presentation, diagnosis, treatment. *Experimental oncology*, (35,№ 4), 303-310.

[2]. Statistical analysis of breast cancer in India, (2019). Available at: https://www.biospectrumindia.com/views/21/15300/statistical-analysis-of-breast-cancer-in-india.html.

[3]. Malvia, S., Bagadi, S. A., Dubey, U. S., & Saxena, S. (2017). Epidemiology of breast cancer in Indian women. *Asia-Pacific Journal of Clinical Oncology*, *13*(4), 289-295.

[4]. Mehra, R. (2018). Breast cancer histology images classification: Training from scratch or transfer learning?. *ICT Express*, *4*(4), 247-254.

[5]. Senapati, M. R., Panda, G., & Dash, P. K. (2014). Hybrid approach using KPSO and RLS for RBFNN design for breast cancer detection. *Neural Computing and Applications*, *24*(3-4), 745-753.

[6]. Acharya, U. R., Ng, E. Y. K., Tan, J. H., & Sree, S. V. (2012). Thermography based breast cancer detection using texture features and support vector machine. *Journal of medical systems*, *36*(3), 1503-1510.

[7]. Bardou, D., Zhang, K., & Ahmad, S. M. (2018). Classification of breast cancer based on histology images using convolutional neural networks. *IEEE Access*, *6*, 24680-24693.

[8]. Lin, C. M., Hou, Y. L., Chen, T. Y., & Chen, K. H. (2013). Breast nodules computer-aided diagnostic system design using fuzzy cerebellar model neural networks. *IEEE Transactions on Fuzzy Systems*, *22*(3), 693-699.

[9]. Ahmad, F., Isa, N. A. M., Hussain, Z., Osman, M. K., & Sulaiman, S. N. (2015). A GA-based feature selection and parameter optimization of an ANN in diagnosing breast cancer. *Pattern Analysis and Applications*, *18*(4), 861-870.

[10]. Karabatak, M. (2015). A new classifier for breast cancer detection based on Naïve Bayesian. *Measurement*, *72*, 32-36.

[11]. Verma, K., Singh, B. K., & Thoke, A. S. (2015). An enhancement in adaptive median filter for edge preservation. *Procedia Computer Science*, *48*(C), 29-36.

Cite this article as: Surendra Prasad M and Manimurugan. A New Modified Recurrent Extreme Learning with PSO Machine Based on Feature Fusion with CNN Deep Features for Breast Cancer Detection. J. Comput. Sci. Intell. Technol. 2020; 1(3): 15–21. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 3, pages: 22–28, December 2020

Face Recognition Framework based on Convolution Neural Network with modified Long Short Term memory Method

[1]**Sushmitha Parikibanda**

[1]*National College of Ireland, Dublin, Ireland.*

[**]***Corresponding Author:*** ***sushmithaparikibanda@gmail.com***

Received: *20.08.2020,*
Revised: *14.11.2020,*
Accepted: *16.12.2020,*
Published: *22.12.2020*

DOI:
10.53409/mnaa.jcsit20201304

Abstract: For real-world applications, such as video monitoring, interaction between human machines and safety systems, face recognition is very critical. Deep learning approaches have demonstrated better results in terms of precision and processing speed in image recognition compared to conventional methods. In comparison to traditional methods. While facial detection problems with different commercial applications have been extensively studied for several decades, they still face problems with many specific scenarios, due to various problems such as severe facial occlusions, very low resolutions, intense lighting and exceptional changes in image or video compression artifacts, etc. The aim of this work is to robustly solve the issues listed above with a facial detection approach called Convolution Neural Network with Long short-term Model (CNN-mLSTM). This method first flattened the original frame, calculating the gradient image with Gaussian filter. The edge detection algorithm Canny-Kirsch Method will then be used to identify edge of the human face. The experimental findings suggest that the technique proposed exceeds the current modern methods of face detection.

Keywords: Brain tumor, Computer-Aided Diagnosis, Deep Learning, Fusion Feature, Recurrent Extreme Learning Machine, Artificial Bee Colony.

I. INTRODUCTION

Face recognition is the mechanism by which the vision system identifies a particular person's face. Due to its use in surveillance systems, access control, video monitoring, business areas and also in social networks such as Facebook it was a key tool for human-computer interaction. Face recognition has once again gained attention since the rapid growth of artificial intelligence due to its non-intrusive nature and because it is the predominant way to recognize individuals by contrasting them with other forms of biometric techniques. Without the awareness of the subject individual in an unregulated setting, face recognition can be easily verified. When investigating the history of face recognition, several academic papers have studied it[1]. Traditional approaches focused on superficial learning face challenges such as pose variation, facets, scene lighting, image context ambiguity, and changes in facial expression as in references [2]. Shallow learning methods use only certain basic image features to extract sample functions and rely on artificial experience.

Deep learning will extract more complex facial characteristics [3]. Deep learning makes significant strides in solving problems that have for many years been limiting the best efforts of the artificial intelligence community. It has proven excellent by exposing high-dimensional data tough structures and thus applies in many fields of science, industry and government. It tackles the question of learning hierarchical representation with one or more algorithms and has predominantly broken records in the fields of image recognition, natural language processing, semantic segmentation and a variety of other scenarios in the real world [5].

Deep learning methods like CNN, the Stacked Autoencoder and the Deep Belief Network (DBN) are distinct. In image recognition and facial recognition, CNN is also used. CNN is a type of artificial neural networks which use the convergence methodology to extract the input data features to increase features. Furthermore, the use of computer vision approaches has increased through CNN's use in the resolution of

many other computer vision activities, such as object detection and identification, segmentation, optical character recognition , facial expression analyzes, age estimates, and so on.

The effect of that is to reduce the memory needs and to reduce correspondingly the number of parameters to be learned. This increases the efficiency of the algorithm. At the same time, the images must be preprocessed or derived from other computer algorithms. Such operations, however, are rarely required for CNN processing images. It's another algorithm that can't learn by computer. In-depth research still exists several limitations. For this purpose, Convolutional Neural Network with Long Short Term (CNN-mLSTM) changed memory was proposed. The original image was initially smoothed with a Gaussian filter during the preprocessation step, and its gradient value was measured. The edge detection algorithm Canny-Kirsch Method is then used to detect edge of the human face. After this function is extracted using a self-Residual Attention-based Network (SRANet), the device collects the global dependencies of spatial and channel dimensions for discriminatory facial feature integration.

The structuring of the rest of paper is as follows: In Section 2, discuss existing face recognition methods. Section 3 describes the proposed methodology. Section 4 describes the results and discussion; in Section 5, conclude the paper and define the future work.

II. RELATED WORK

An inherent correlation between detection and augmentation is proposed under the deep cascading multifunction system in [6] to improve their performance. We exploit a cascaded architecture with three stages in which deep convolutional neural networks are deliberately created to roughly forecast the face and landmark.

A CNN cascade multi-task system [7] that contains both tasks. We show that multi-task face recognition learning and the head pose estimation helps obtain more representative characteristics.

For face detection, [8] is proposed based on skin color segmentation and facial properties. An Algorithm is a competent analytical tool to evaluate different color patterns such as RGB, YcbCr, and HSV, as well as their skin color detection combinations.

In [9] a new approach called the algorithm DP-Adaboost proposed for the detection of the human face and the improvement of the right rate of detection. The multi-angle face is identified by an enhanced Adaboost algorithm with the combination of a frontal face grader and a profile face grader.

The architecture in cascading with three phases carefully designed [10] to forecast the presence of faces, based on deep convolutional networks. In [11] an important facial recognition system has been developed to index a specific face from various video images. The classifier used is the enhanced neural grid that optimizes weight factors with the modified cuckoo search algorithm.

III. PROPOSED METHODOLOGY

This research proposes a facial recognition system with modified Long Short Term (CNN-mLSTM) memory from the Convolutional Neuronal Network. The architectural structure is of three layers and recognizes all picture regions containing images. The face detection is an automated face reconnaissance device pre-processing level. In the first step, Retinex adaptive filter is used for improving the image to eliminate unnecessary noise. At the next level, a Canny-Kirsch Method edge detection is performed on the face edge and functional extraction takes place using SCA and SSA. The CNN-mLSTM then classifies the unknown class or class that is not face to face as either face to face or non-face. The block diagram of the proposed CNN-mLSTM based face recognition is illustrated in Figure 1.

3.1. Input database

The database was established in February 2002 at the IIT Kanpur campus. There are forty subjects, and 11 different images per subject. Few more photos are included for certain subjects. All pictures are brightly homogeneous at the backdrop. The subjects are frontally straight. The pictures had been in JPEG format. Each image has a resolution of 640x480, with 256 gray per pixel. The male and female images have been put in two principal folders. Every subject has eleven distinct images in both folders. Database are variations based on orientation and emotion.

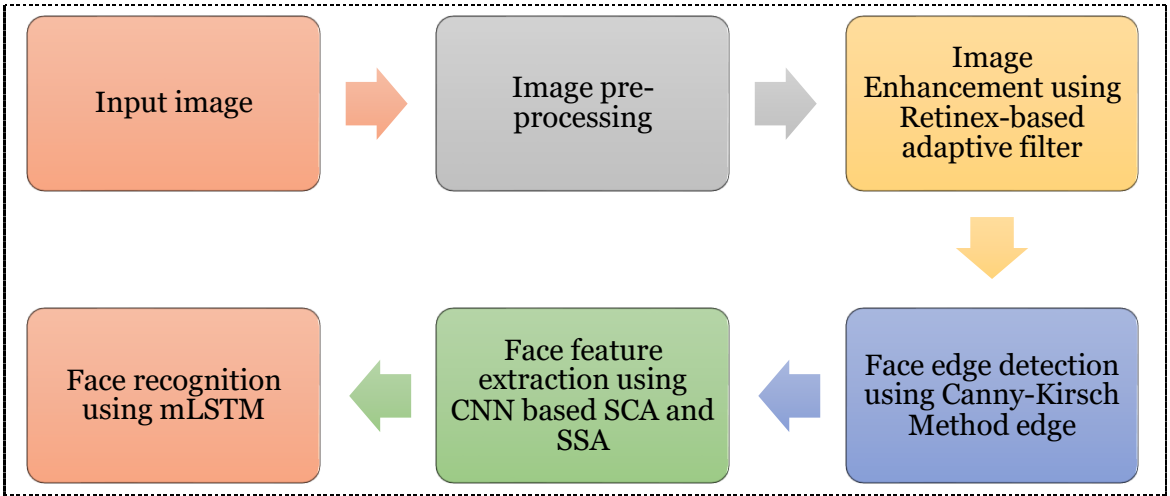

Figure 1. Flow diagram of Proposed Face detection.

3.2. Image Pre-processing Step

In pre-processing of images, since the various conditions, such as the different sources, can affect the final result. The proposed system would change the facial expression luminous intensity to be registered in the picture files. Higher pixel interval brightness to calculate the mean in this image, and used this average value as a reference and calculation in equation (1 & 2);

$$R_a = \frac{\sum_1^n \max N_R}{n}, G_a = \frac{\sum_1^n \max N_G}{n}, B_a = \frac{\sum_1^n \max N_B}{n} \quad (1)$$

$$R' = \frac{255}{R_a} * N_R, G' = \frac{255}{G_a} * N_G, B' = \frac{255}{B_a} * N_B \quad (2)$$

Adjustments of the original picture pixel based on mean values. Where N_R, N_G and N_B represent pixel values of the original image. In the color channel interval of n, R, G and B represent the average value of a pixel, and here n represent the total number of pixels to obtain a brightness range, and represent pixel values after the configuration. This pre-processing image is used to change the low light images.

3.3. Image Enhancement using Retinex-based adaptive filter

Retinex based adaptive filter is a proposed framework for enhancing color images in this segment. This system can be used to improve conventional 24-bit images as well as to compress high dynamic range images generated from raw format or multiple exposure techniques that are linear RGB images. Let Retinex theory describe the portion of luminance treated as:

$$Retinex_Y = \log_{10}(I'_Y) - \log_{10}(mask) \quad (3)$$

Where $\log_{10}(I'_Y)$ is the Y portion of the non-linear RGB image I and translated to the color space of YcbCr. The last term mask, is a matrix representing the weighted average of its surrounding for each pixel. A significant factor is how you describe this surrounding and its corresponding weights. A traditional approach is to define the mask with a filter conveying the color.

$$mask = I'_Y * \mathrm{LP}F \quad (4)$$

Where LPF is a circularly symmetric low-pass filter that is strictly defined by a 1-dimensional function rotated around the z axis and the 1-dimensional curve is typically defined by a simple Gauussian or Gauussian function composition. The 1-dimensional radial function is a Gaussian curve, whose spatial constant varies depending on the local contrast of the face image. The spatial constant is given by equation (5) with an initial value ÿ. If a strongly contrasted edge is marked along the radius, there will be 8 divided into σ.

$$\sigma = \frac{r_{max}}{8}, where\ \max(I_{size}) = r_{max} \quad (5)$$

Although the weights and support of the filter are modified for each pixel, the mask is determined sequentially pixel after pixel and mask(x, y) is the weighted sum of elements in the Coordinate(x, y) pixel surround.

$$mask(x,y) = \int_{\theta=0}^{360} \int_{r=0}^{r_{max}} I_Y(x + \cos\theta), y + \sin(\theta) e^{-\frac{r^2}{\sigma^2}} \quad (6)$$

Where σ is the Gaussian spatial constant which varies in the direction of radius. In this way, the support of the filter follows essentially the high contrast facial edges of the image. Using the Kirsch edge detector these face edges are detected.

3.4. *Face Feature Extraction and recognition using CNN*

A face recognition method based on CNN-CEC-LSTM is proposed in this section. And the network used here consists of nine layers. These layers contains convolution layers, pooling layers, full-connected layers and Softmax regression layer. The convolution layers and the pooling layers are used for feature extraction followed by full-connected layers, and the last layer uses a CEC-LSTM classifier with strong non-linear classification capability. The CNN based feature extraction is illustrated in Figure 4.

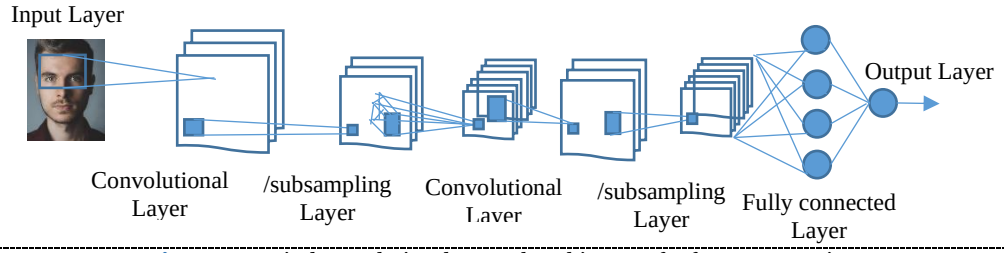

Figure 1. Typical convolutional network architecture for feature extraction

CNN comprises three types of layers: input, convolution, and pooling, and layers which are fully connected. In the input layer, the obtained information consists of multiple image sequences $\{I_1, I_2, \ldots, I_n\}$, which consist of the dataset indicator diagram. Built on the original CNN architectures, add attention modules to each of the ResNet structure's residual bottlenecks to get a refined face function. In particular, the proposed attention module consists of two blocks called the self-residual channel attention module and the self-residual spatial attention module, which sequentially learns the channel relationship matrix and spatial relationship matrix, and then achieves the refined function by multiplication of matrixes.

For instance, assumed an intermediate feature map FM, the channel refined feature FC and the spatially refined feature FS can be sequentially obtained. Moreover, it argues that the features derived from the global average pooling layer are not adequately inclusive for deep face recognition, so instead, use a completely linked layer. With the above modifications, the information redundancy between channels can be reduced as well as the most important part of face images can be learnt. At last, the refined function can be obtained by residual shortcut learning. The function vectors received will then be fed to the sequential sheet. To capture long distance dependency, mLSTM for vector composition is inserted into the sequential layer.

3.5. *Modified Long Short Term Memory (mLSTM) for Face Recognition*

Recurrent Neural Network (RNN) is an enhanced version of LSTM. Instead of traditional simple RNN modules, LSTM implements memory blocks to tackle the issue of the gradient vanishing and bursting. LSTMs are also better able to manage long term dependencies than conventional RNNs. This means that LSTMs can recall and relate past knowledge to the present (which actually lags very far back in time as opposed to the present). A memory block in LSTM is a complex processing unit, which consists of one or more memory cells. A pair of multiplicative gates are being used as gateway for input and output. A collection of adaptive, multiplicative gates regulates the full operations of a memory block. The output of the input gate makes or discards activity for the input flow to a memory cell from a cell activation.

The efficiency of the output gate makes or discards operation to other nodes for an output state of a memory cell. Forgetting gate and peephole connections were integrated into the current LSTM network as research on LSTM progressed. Instead of the Persistent error carousel (CEC), the Forget

gate is used. The forget gate allows to forget or reset a memory cell's states. The peephole connections are made to all of its gates from a memory cell. They study both the exact timing of the outputs and the internal state of a memory cell. The mLSTM function is as follows.

CNN's features input sequence is fed into the mLSTM architecture. In the recurrent hidden layer (h) of LSTM architecture, the output sequence of continuous write, read, and reset operations by three multiplicative units (input (i), output (o), and forget gate (f)) on the memory cell (c) is calculated iteratively from $j = 1,2, \dots, j+1$. The sequence of operations taking place in mLSTMs at time step j can be fleetingly signified by the below equation (7).

$$i_t = \sigma\left(w_{xi}x_j + w_{hi}h_{j-1} + w_{ci}c_{j-1} + b_i\right) \quad (8)$$

$$f_j = \sigma\left(w_{xf}x_j + w_{cf}h_{j-1} + w_{cf}c_{j-1} + b_f\right)$$

$$c_j = f_j \odot c_{j-1} + i_j \odot \tan h\left(w_{xc}x_j + w_{hc}h_{j-1} + b_c\right)$$

$$o_j = \sigma\left(w_{xo}x_j + w_{co}h_{j-1} + w_{co}c_{j-1} + b_o\right)$$

$$h_j = o_j \odot \tanh\left(c_j\right)$$

$$y_j = w_{yh}h_j + b_y$$

Where ⊙ remains the scalar product of two vectors and $\sigma()$ means the standard logistics sigmoid function defined as follows (9)

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (9)$$

At this time weight matrices denoted as w and bias vectors b are used to build connections bet ween the input layer, output layer and memory block. CNN in this CNN-mLSTM consists only of layer convolution and layer maxpooling. The maxpooling layer output is transmitted to the subsequent LSTM layer.

$$y_j = CNN(x_i) \quad (10)$$

Here x_i remains the initial input vector to the CNN network with the class label and y_j stands the output of the CNN network to be nourished to the next mLSTM network x_i the feature vector fashioned from the max-pooling operation in CNN. Learning about the long-range temporal dependencies is fed to the mLSTM.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed CNN-mLSTM is evaluated in this section, and the performance results are compared with existing CNN [7], DP-Adaboost [9] and improved neural network [11] face detection schemes. The facts given below show that the device proposed has achieved better performance in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (6):

$$Precision = \frac{TP}{FP+TP} \quad (11)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (12):

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

F-measure: this is also referred to as F 1-score, and as in equation (13) is the harmonic mean of precision and recall:

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (13)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (14):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (14)$$

Where true positive (TP) samples are properly classified faces, false positive (FP) samples are incorrectly classified correctly, True negative (TN) samples are not properly classified, and false negatives (FN) are incorrectly classified faces.

4.1. Precision Rate comparison

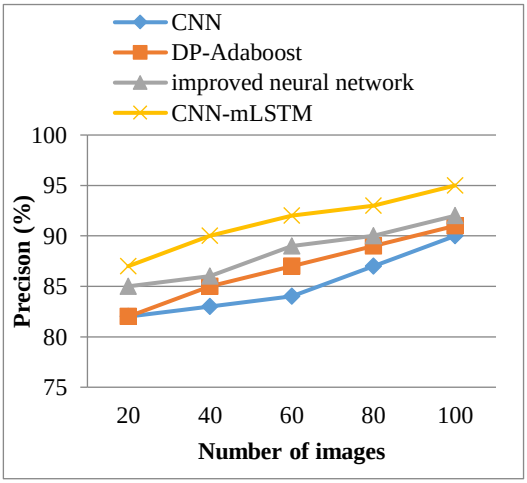

Figure 4. Representation of Precision Comparison

From the above Figure 4, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as CNN, DP-Adaboost, improved neural network and CNN-mLSTM. When the number of records increases according to the precision value. From this graph, it is learned that

the proposed CNN-mLSTM offers 95% higher precision than previous methods that yield better results in the detection of face. The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

No.of images	CNN	DP-Adaboost	improved neural network	CNN-mLSTM
20	82	82	85	87
40	83	85	86	90
60	84	87	89	92
80	87	89	90	93
100	90	91	92	95

4.2. Recall comparison

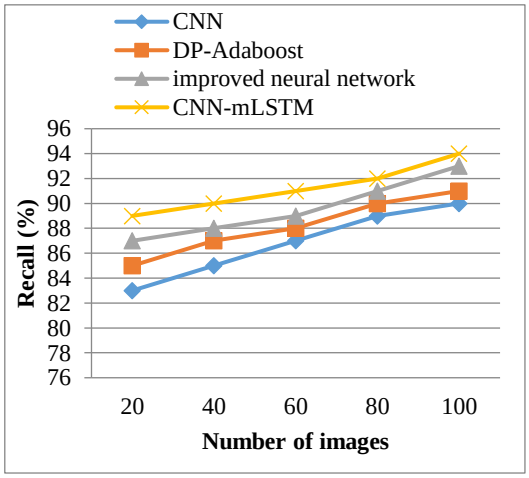

Figure 5. Representation of Recall Comparison

From the above Figure 5 the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as CNN, DP-Adaboost, improved neural network and CNN-mLSTM. Increasing the number of images often increases the correct value for the recall. Through this graph, it is discovered that the current CNN-mLSTM offers recall 94% higher than previous methods. The explanation for this is that the CNN-mLSTMextracts the features directly which will enhance the face detection results. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

No.of images	CNN	DP-Adaboost	improved neural network	CNN-mLSTM
20	83	85	87	89
40	85	87	88	90
60	87	88	89	91
80	89	90	91	92
100	90	91	93	94

4.3. F-measure Rate comparison

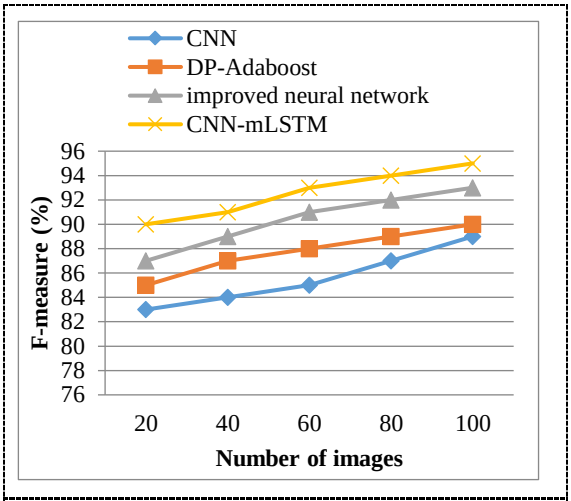

Figure 6. Representation of F-measure Comparison

From the above Figure 6, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as CNN, DP-Adaboost, improved neural network and CNN-mLSTM. When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed CNN-mLSTM offers 95% higher f-measure than previous methods. Therefore the proposed CNN-mLSTM algorithm is stronger than the current algorithms in terms of better performance of classifying face. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

No.of images	CNN	DP-Adaboost	improved neural network	CNN-mLSTM
20	83	85	87	90
40	84	87	89	91
60	85	88	91	93
80	87	89	92	94
100	89	90	93	95

4.4. Accuracy comparison

From the above Figure 7 the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as CNN, DP-Adaboost, improved neural network and CNN-mLSTM. From this graph it is known that the proposed CNN-mLSTM algorithm is higher than the existing algorithms with a high precision rate of 96% in terms of better template matching results.

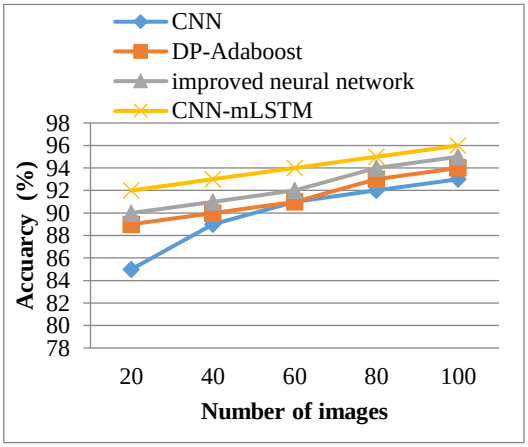

Figure 7. Representation of Accuracy Comparison

This is due to the automatic extraction of the features using CNN in the CNN-mLSTM classification algorithm, which increases the classification precision resulting in Brain tumour. The numerical results of Accuracy Comparison is shown in Table 4.

Table 4. The numerical results of Accuracy Comparison

No.of images	CNN	DP-Adaboost	improved neural network	CNN-mLSTM
20	85	89	90	92
40	89	90	91	93
60	91	91	92	94
80	92	93	94	95
100	93	94	95	96

V. CONCLUSION AND FUTURE WORK

In this work, face recognition is done using proposed CNN-mLSTM. The overall performances were obtained using the different number of training images and test images. Initially, after the pre-processing the Retinex-based adaptive filter is applied to enhance the face images. The convolutional neural networks achieve the best results of feature extraction so far. This work proposed to use a LSTM network and compare its performance with a standard MLP network for face classification problems. The mLSTM network presented for face recognition which can attain better performance in terms of correct classification rates in all the three proposed face classification tasks, showing that it is a powerful tool in face recognition applications, even if dealing with a reduced training set. The CNN-mLSTM achieves a much better recognition performance than the conventional schemes such as CNN, DP-Adaboost and improved neural network. The proposed system can be extended by using the classifiers like deep learning with the various optimization schemes like, genetic algorithm etc.

REFERENCES

[1]. Ding, C., & Tao, D. (2016). A comprehensive survey on pose-invariant face recognition. *ACM Transactions on intelligent systems and technology (TIST)*, *7*(3), 1-42.

[2]. Lei, Z., Yi, D., & Li, S. Z. (2015). Learning stacked image descriptor for face recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, *26*(9), 1685-1696.

[3]. Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2015). Learning deep representation for face alignment with auxiliary attributes. *IEEE transactions on pattern analysis and machine intelligence*, *38*(5), 918-930.

[4]. Bharati, A., Singh, R., Vatsa, M., & Bowyer, K. W. (2016). Detecting facial retouching using supervised deep learning. *IEEE Transactions on Information Forensics and Security*, *11*(9), 1903-1913.

[5]. Ding, C., & Tao, D. (2016). A comprehensive survey on pose-invariant face recognition. *ACM Transactions on intelligent systems and technology (TIST)*, *7*(3), 1-42.

[6]. Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, *23*(10), 1499-1503.

[7]. Wu, H., Zhang, K., & Tian, G. (2018). Simultaneous face detection and pose estimation using convolutional neural network cascade. *IEEE Access*, *6*, 49563-49575.

[8]. Yadav, S., & Nain, N. (2016). A novel approach for face detection using hybrid skin color model. *Journal of Reliable Intelligent Environments*, *2*(3), 145-158.

[9]. Zheng, Y. Y., & Yao, J. (2015). Multi-angle face detection based on DP-Adaboost. *International Journal of Automation and Computing*, *12*(4), 421-431.

[10]. Luo, D., Wen, G., Li, D., Hu, Y., & Huan, E. (2018). Deep-learning-based face detection using iterative bounding-box regression. *Multimedia Tools and Applications*, *77*(19), 24663-24680.

[11]. Yoganand, A. V., & Kavida, A. C. (2018). Face detection approach from video with the aid of KPCM and improved neural network classifier. *Multimedia Tools and Applications*, *77*(24), 31763-31785.

Cite this article as: Sushmitha Parikibanda. Face Recognition Framework based on Convolution Neural Network with modified Long Short Term memory Method. J. Comput. Sci. Intell. Technol. 2020; 1(3): 22–28. ©JCSIT, MNAA PUB WORLD, 2020.

Journal of Computational Science & Intelligent Technologies, eISSN: 2582-9041

Volume 1, Issue 3, pages: 29–34, December 2020

Performance Analysis of Deep Belief Neural Network for Brain Tumor Classification

[1]Sreenivas Eeshwaroju & [2]Praveena Jakula

[1] *Harman Connected Services, Novi, Michigan, USA.*
[2] *Intel Corporation, USA.*

[**]*Corresponding Author: Sreenivas.Eeshwaroju@harman.com*

Received: *30.08.2020,*
Revised: *20.11.2020,*
Accepted: *17.12.2020,*
Published: *22.12.2020*

DOI:
10.53409/mnaa.jcsit20201305

Abstract: The brain tumors are by far the most severe and violent disease, contributing to the highest degree of a very low life expectancy. Therefore, recovery preparation is a crucial step in improving patient quality of life. In general , different imaging techniques such as computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound imaging have been used to examine the tumor in the brain, lung , liver, breast , prostate ... etc. MRI images are especially used in this research to diagnose tumor within the brain with classification results. The massive amount of data produced by the MRI scan, therefore, destroys the manual classification of tumor vs. non-tumor in a given period. However for a limited number of images, it is presented with some constraint that is precise quantitative measurements. Consequently, a trustworthy and automated classification scheme is important for preventing human death rates. The automatic classification of brain tumors is a very challenging task in broad spatial and structural heterogeneity of the surrounding brain tumor area. Automatic brain tumor identification is suggested in this research by the use of the classification with Deep Belief Network (DBN). Experimental results show that the DBN archive rate with low complexity seems to be 97 % accurate compared to all other state of the art methods.

Keywords: Magnetic Resonance Imaging, Brain tumor and Deep Belief Network .

I. INTRODUCTION

The incidence of central nervous system (CNS) tumors in India ranges from 5 to 10 per 100,000 population with an increasing trend and accounts for 2% of malignancies [1]. Brain MRI image is mainly used to identify the process of modeling tumor and tumor progression. Such information is used mainly for procedures of tumor diagnosis and treatment. The MRI image contains more detail about the medical condition provided than the CT or ultrasound condition. The MRI image contains accurate brain structure and irregularity identification evidence in brain tissue.

In addition, from the time when it became possible to scan and bring medical imagesto the machine, Scholars provided unlike automated methods for brain tumor detection and typecataloging using brain MRI imaging. In comparison, Neural Networks (NN) and Support Vector Machine (SVM) have been the most widely used approaches for their successful implementation over the last few years[2].

Even so recently, models of Deep Learning (DL) set a compelling trend in machine learning as the massive underground architecture can effectively represent complex relationships without needing a large number of nodes, such as in the superficial architectures. K-Nearest Neighbor (KNN), and Vector Machine Support (SVM). As a result, they evolved rapidly to become the state of the art in areas such as medical image processing, medical informatics and bioinformatics, apart from in health informatics.

This work proposes the DBN for the classification of brain tumors in MRI with this motivation. The proposed model to improve tumor and non-tumor detection rates with high accuracy. The primary contributions of this work are as follows:

- It is the first systematic method for the classification of MRI brain tumors using new Deep Neural Belief (DBN) networks
- Evaluation of the performance of the new framework demonstrating state-of-the-

art output in contrast with current approaches.

The rest of the paper is organised as follows. Section 2 deals with the classification of MRI brain tumor. Section 3 describes the methods proposed for defining the MRI brain tumor classification with DBN. Section 4 discusses the experimental findings. Section 5 includes the conclusion and prospective work.

II. RELATED WORK

In a framework combined with a global probabilistic image model, local tissue intensity model and priors of the Markov Random Field[2] are introduced which may impede the automated methods of tissue classification in Brain IRM. A different method for the classification of MRI brain image is suggested by that of the integration of wavelet entropy-based web plots and probabilistic neural network. The two-stage classification process utilizes Spider Web Plots based on wavelets for the extraction of the characteristics and a probabilistic neural network for the classification[3].

The [4] fuzzy logic hybrid kernel has been created and applied for the automatic classification of four cancer types, including the meningioma, glioma, astrocytoma, and metastases called fuzzy logic hybrid kernel SVM, for the support of the Vector Machine.

In [5] a two techniques fusion is employed to identify the imperatives in brain tumor, namely the Tolerance Rugh Set (TRS) and the FireflyAlgorithm (FA). With [6] a CAD system can be connected to a SVM classification using a quadratic kernel function to support radiologists in their diagnosis procedures through a series of pre-processing , segmentation and extraction steps.

The self-organizing neural mapping network initiates [7] by training the features extracted from the discrete wavelet mixing wavelets and thus train the K-nearest neighbor in the filter factors and complete the testing process in two phases.

In [8], a brain tumor segmentation and classification method for MRI scans was proposed. The integrated characteristics are subsequently provided in five classes to the random forest classifier. DBN is being used to classify brain tumor images of MRI in this proposed system.

III. PROPOSED METHODOLOGY

This section presents the overall structure of the classification system for MRI brain tumors, and the architecture is shown in Figure 1. The analysis and modeling is made up ofthree stages, namely pre-processing, extraction of features, and classification. The training brain MRI image is provided as input to the system during the training process and is subjected to all of the above listed steps. A classifier model is selected DBN, where the extracted features are used to train the classifier along with the class labels. The DBN classifier now recognizes the features of the test image and will allocate a class label to the test image as either 'tumor affected' or 'tumor unaffected' with the information it has already acquired during the training process.

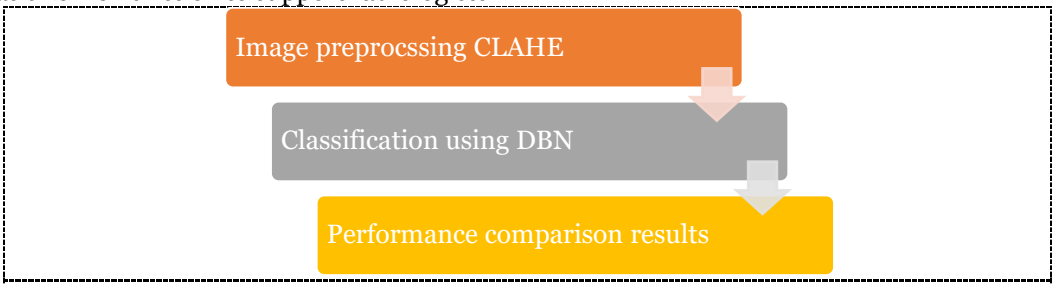

Figure 1. The architecture diagram of classification of MRI using DBN.

3.1. *Dataset Collection and Description*

The data collection of brain tumors obtained from [7] contains 30,064 T1-weighted contrast-enhanced images. The dataset is compiled from 233 patients with three brain tumor types: meningioma (708 images), glioma (1426 images), and pituitary tumor (930 images). What file reserves a struct which comprises different image fields. The label (1 for meningioma, 2 for glioma, 3 for pituitary tumor), PID (patient ID), image info, tumor border, and tumor mask are included.

3.2. *Image Preprocessing using CLAHE*

The first phase in the method for classifying brain tumors in MRI is preprocessing. It preprocesses the input image in such a way that the output image shifts. For this work wiener filter Contrast Limited Adaptive Histogram Equalization (CLAHE) does not find the image in its entirety but separates it as tiles and improves the contrast of

each tile while enhancing the contrast of the whole image. Used for reducing noise effects in the image. The output is taken as the input to the CLAHE after applying wiener filter which is used to enhance the contrast of the given input image. CLAHE algorithm is prearranged as follows:
Step 1: Read the input image.
Step 2: Put on Wiener filter to remove noise.
Step 3: Look after discover the number of frequencies for each pixel.
Step4: The transformation function is calculated using probability density of the input MRI brain image grayscale significance is utilized for amend all histograms, where n remains the total number of pixels in the input MRI brain image and n_j stands the input pixel number of grayscale value j.
Step 5: The grayscale values for the MRI-image are accustomed established on the effects of modified histograms besides bilinear interpolation actuality used modify the neighboring MRI-image.
Step 6: The perseverance of histogram equalization mapping is to deliver input MRI brain image intensity values in such a way as to deliver an essentially uniform distribution of the histogram from the resulting images. The histogram of MRI brain image with gray levels in the assortment $[0, L-1]$ is through discrete function $p(gray_k) = \frac{n \times k}{n}$ somewhere $gray_k$ remains the kth gray level, $n \times k$ stands the number of pixels in MRI brain image with the gray level, here $k = 0, 1, 2, \ldots, L-1$. Mainly, $p(gray_k)$ provides an evaluation of the occurrence probability of gray level in MRI brain image.
Step 7: Display the enhanced MRI brain image.

3.3. MRI brain tumor classification Using Deep Belief Networks

Professor Geoffrey Hinton develops Deep Belief Networks (DBN) consisting of two distinct forms of neural networks-Belief Networks and Restricted Boltzmann Machines-to address the weakness of earlier neural networks. Here the emphasis was on the DBN-based improved Restricted Boltzmann Machines.

Boltzmann System is a recurrent stochastic neural network with binary stochastic units and undirected edges between units. Unfortunately, Boltzmann machine learning is inefficient, and has a problem of scalability. This resulted in the implementation of Restricted Boltzmann Machine (RBM)[10], which has one layer of hidden units and limits relations between hidden units. It allows for more effective learning algorithms[11]. The RBM structure is shown in figure 2 below.

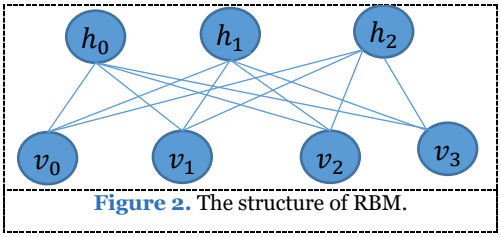

Figure 2. The structure of RBM.

As Deep Belief Networks (DBN) name indicates, it As the name of Deep Belief Networks (DBN) indicates, it is networks of multilayer beliefs[12]. That layer is Restricted Boltzmann Machine and are stacked to create DBN. The first step of training DBN is to use the Gaussian Bernoulli RBM algorithm to learn a layer of features from the visible units 8. Then the next step is to treat the activations of features previously trained as visible units and learn features in a second hidden layer.
Eventually, when learning for the final secret layer is reached the entire DBN is trained. The Gaussian Bernoulli RBM is used, since the input data of the deep architecture are real values. It is composed of a visible layer, a hidden layer and a layer of output. The energy relationship can be written in equation (1-3) and the conditional probability distribution of:

$$E(v, h|\theta) = \sum_{i=1}^{V} \frac{(v_i - a_i)^2}{2\sigma_i^2} - \sum_{i=1}^{V} b_j h_j - \sum_{i=1}^{V} \sum_{j=1}^{H} \frac{v_i}{\sigma_i} h_j w_{ij} \quad (1)$$

$$p(h_i|v;\theta) = \delta\left(\sum_{i=1}^{V} w_{ij} v_i + b_j\right) \quad (2)$$

$$p(v_i|h;\theta) = N\left(\sigma_i \sum_{j=1}^{H} w_{ij} h_j + a_i, \sigma_i^2\right) \quad (3)$$

where $\theta = (w, a, b)$, w_{ij} is the connection weight between the visible unit v_i and hidden unit h_j, a_i is the bias value of v_i, b_i is the bias value of h_j, $\delta(x)$ can be logistic function and $N(\mu, \sigma_i^2)$ is the probability with the mean μ and variance σ^2. The flowchart of DBNs depth is given in Figure 3.

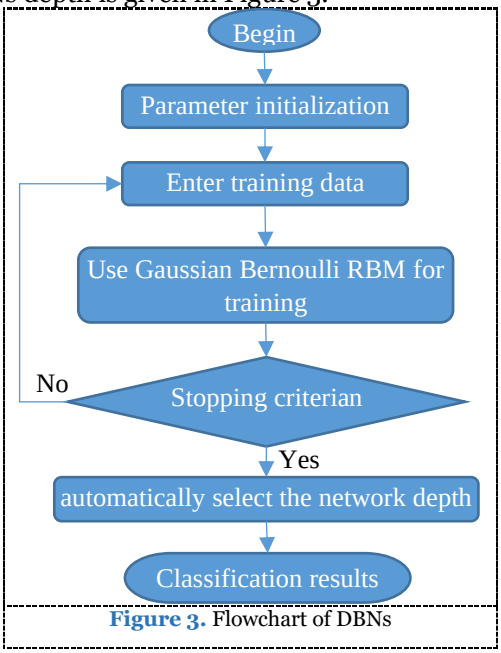

Figure 3. Flowchart of DBNs

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The performance of the proposed DBN is evaluated in this section, and the performance results are compared with existing probabilistic neural network [3], Fuzzy Logic-Based Hybrid Kernel SVM [4] and KNN [7] schemes. The performance measurement is done in terms of precision, f-measurement, recall and accuracy.

Precision: It reflects the proportion of positive samples correctly classified as expected in equation (5):

$$Precision = \frac{TP}{FP+TP} \quad (5)$$

Recall: The recall of a classifier reflects the positive samples properly assigned to the total number of positive samples and is calculated as in equation (6):

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

F-measure: this is also referred to as F 1-score, and as in equation (7) is the harmonic mean of precision and recall:

$$F - measure = \frac{2*(Recall * Precision)}{(Recall + Precision)} \quad (7)$$

Accuracy: This is one of the most frequently used performance classification measures and is defined as a ratio between the correctly classified samples and the total number of samples as in equation (8):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Where true positive (TP) samples are properly classified as no tumor, false positive (FP) samples are incorrectly classified as tumor, True negative (TN) samples are properly classified as tumor, and false negatives (FN) are incorrectly classified as tumor.

4.1. Precision Rate comparison

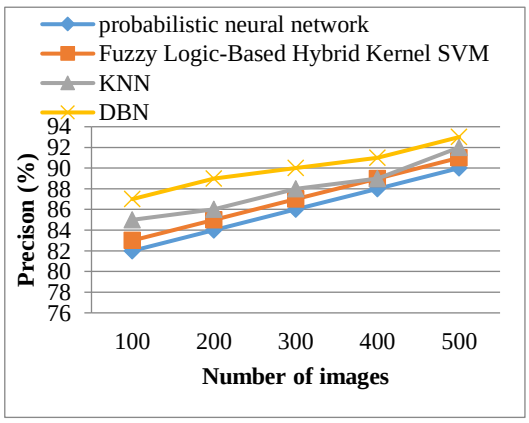

Figure 4. Representation of Precision Comparison

From the above Figure 4, the graph shows how accurate the number of images in the specified datasets is compared. These methods are implemented as probabilistic neural network, Fuzzy Logic-Based Hybrid Kernel SVM, KNN and DBN. When the number of records increases according to the precision value, from this graph, it is learned that the proposed DBN offers 93% higher precision than previous methods that yield better results in the classification of MRI brain tumor due to Gaussian Bernoulli RBM. The numerical results of Precision Comparison is shown in Table 1.

Table 1. The numerical results of Precision Comparison

No.of images	probabilistic neural network	Fuzzy Logic-Based Hybrid Kernel SVM	KNN	DBN
100	82	83	85	87
200	84	85	86	89
300	86	87	88	90
400	88	89	89	91
500	90	91	92	93

4.2. Recall comparison

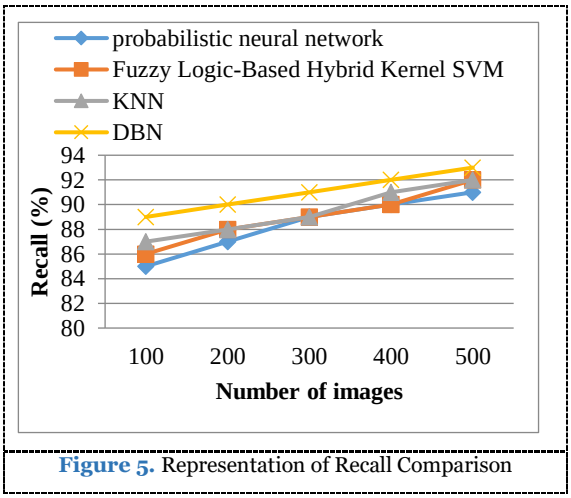

Figure 5. Representation of Recall Comparison

From the above Figure 5, the graph illustrates the recall relation for the number of images in the listed datasets. These methods are implemented as Fuzzy Logic-Based Hybrid Kernel SVM, KNN and DBN. Increasing the number of images often increases the correct value for the recall. Through this graph, it is discovered that the current DBN offers recall 93% higher than previous methods. The explanation for this is that the DBN extracts the features directly, which will enhance the detection and classification of brain tumor. The numerical results of Recall Comparison is shown in Table 2.

Table 2. The numerical results of Recall Comparison

No.of images	probabilistic neural network	Fuzzy Logic-Based Hybrid Kernel SVM	KNN	DBN
100	85	86	87	89
200	87	88	88	90
300	89	89	89	91
400	90	90	91	92
500	91	92	92	93

4.3. F-measure Rate comparison

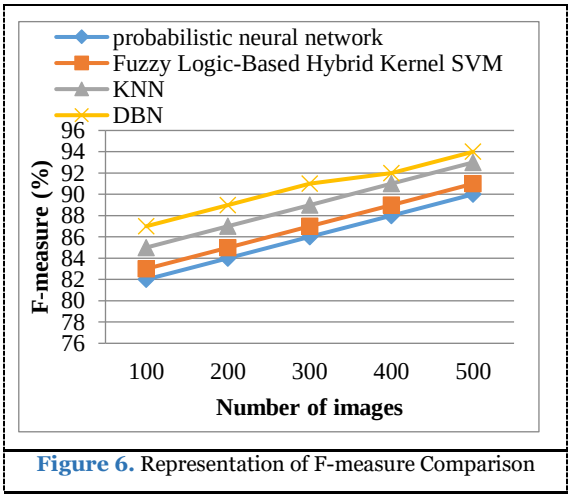

Figure 6. Representation of F-measure Comparison

From the above Figure 6, the graph explains the f-measure relation for the number of images in the given datasets. These methods are implemented as Fuzzy Logic-Based Hybrid Kernel SVM, KNN and DBN. When the number of data is increased, and the f-measure value is increased accordingly. From this graph it is learned that the proposed DBN offers 94% higher f-measurement than previous methods. Therefore the proposed DBN algorithm is stronger than the current algorithms in terms of better performance of classifying MRI brain tumor. The numerical results of F-measure Comparison is shown in Table 3.

Table 3. The numerical results of F-measure Comparison

No.of images	probabilistic neural network	Fuzzy Logic-Based Hybrid Kernel SVM	KNN	DBN
100	82	83	85	87
200	84	85	87	89
300	86	87	89	91
400	88	89	91	92
500	90	91	93	94

4.4. Accuracy comparison

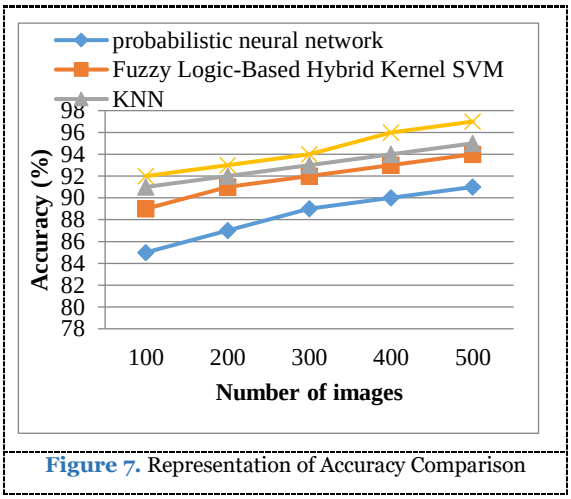

Figure 7. Representation of Accuracy Comparison

From the above Figure 7, the diagram illustrates the processing time relation for the number of images in the specified datasets. These methods are implemented as Fuzzy Logic-Based Hybrid Kernel SVM, KNN and DBN. From this graph, it is known that the proposed DBN algorithm is higher than the existing algorithms with a high precision rate of 97% in terms of better template matching results. This is due to the automatic extraction of the function in the DBN algorithm, which increases the MRI brain tumor classification results. The numerical results of Accuracy Comparison is shown in Table 4.

Table 4. The numerical results of Accuracy Comparison

No.of images	probabilistic neural network	Fuzzy Logic-Based Hybrid Kernel SVM	KNN	DBN
100	85	89	91	92
200	87	91	92	93
300	89	92	93	94
400	90	93	94	96
500	91	94	95	97

V. CONCLUSION AND FUTURE WORK

This research indicates that DBN algorithms are used to extract features for the classification of MRI braiun tumors with a high precision rate of 97%. The machine will considerably divide the tumor into three levels; meningioma, glioma, and pituitary tumor using contrast-enhanced brain MR images of T1 weight. Including more brain MR images with different weights and with various contrast enhancement techniques to allow the architecture to be potentially more versatile and reliable for larger image databases will further increase this architectural grading performance. The proposed model, nevertheless, still poses shortcomings such as long calculation time. The next research material

will be how to refine the algorithm and shorten the run-time.

REFERENCES

[1]. Nair M, Varghese C, Swaminathan R. Cancer: Current Scenario, Intervention Strategies and Projections for 2015. *NCMH Background Papers*; 2015.

[2]. Tohka, J., Dinov, I. D., Shattuck, D. W., & Toga, A. W. (2010). Brain MRI tissue classification based on local Markov random fields. *Magnetic resonance imaging*, *28*(4), 557-573.

[3]. Saritha, M., Joseph, K. P., & Mathew, A. T. (2013). Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. *Pattern Recognition Letters*, *34*(16), 2151-2156.

[4]. Jayachandran, A., & Sundararaj, G. K. (2015). Abnormality segmentation and classification of multi-class brain tumor in MR images using fuzzy logic-based hybrid kernel SVM. *International Journal of Fuzzy Systems*, *17*(3), 434-443.

[5]. Jothi, G. (2016). Hybrid Tolerance Rough Set–Firefly based supervised feature selection for MRI brain tumor image classification. *Applied Soft Computing*, *46*, 639-651.

[6]. Natteshan, N. V. S., & Jothi, J. A. A. (2015). Automatic classification of brain mri images using svm and neural network classifiers. In *Advances in intelligent informatics* (pp. 19-30). Springer, Cham.

[7]. Anitha, V., & Murugavalli, S. J. I. C. V. (2016). Brain tumour classification using two-tier classifier with adaptive segmentation technique. *IET computer vision*, *10*(1), 9-17.

[8]. Usman, K., & Rajpoot, K. (2017). Brain tumor classification from multi-modality MRI using wavelets and machine learning. *Pattern Analysis and Applications*, *20*(3), 871-881.

[9]. J. Cheng, 'brain tumor dataset', 2017. [Online]. Available: https://figshare.com/articles/brain_tumor_dataset/1512427,

[10]. Hinton, G. E. (2012). A practical guide to training restricted Boltzmann machines. In *Neural networks: Tricks of the trade* (pp. 599-619). Springer, Berlin, Heidelberg.

[11]. Hinton, G., Srivastava, N., & Swersky, K. (2012). Neural networks for machine learning. *Coursera, video lectures*, *264*(1).

[12]. Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep machine learning-a new frontier in artificial intelligence research [research frontier]. *IEEE computational intelligence magazine*, *5*(4), 13-18.

Cite this article as: Sreenivas E and Praveena J. Performance Analysis of Deep Belief Neural Network for Brain Tumor Classification. J. Comput. Sci. Intell. Technol. 2020; 1(3): 29–34. ©JCSIT, MNAA PUB WORLD, 2020.