New Intrusion Detection System Based on Neural Networks and Clustering

Kancherla Samata[1*], *Dugyala* Raman[2], *S* Saravanan[1], *R* Saminathan[1]

[1]Department of Computer Science and Engineering, Annamalai University
[2]Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology

Abstract. Efficiency of Intrusion detection systems-IDS are evaluated using parameters like completeness, performance and accuracy. The first important parameter is the completeness, which occurs when the detection of attack fails. This is the most difficult parameter to evaluate compared to the other two parameters. The second one is performance, which indicates the audit events process. When the IDS doesn't work properly or works poorly, the real time detection becomes impossible. Legitimate actions are flagged as anomalous which is termed as inaccuracy. This part needs attention to address the inaccuracies. Optimal solutions must take the inaccuracies into consideration for accuracy, thereby efficiency of IDS. There are different trends in IDS. Some of them are discussed below. Behavior and knowledge-based IDS: Misuse detection, appearance-based detection, behavior detection and anomaly detection etc. There are numerous stability and security issues as a result of the Internet's and computer networks' rapid proliferation. The present study reports the case study of image processing in a fruit grading plant with data safety over cloud with Original Equipment Manufacturer (OEM). How Artificial Neural Networks (ANN) architecture can help is discussed and recommendations are made for impending improvement.

Keywords: Network Artificial Neural Networks, Optimal Intrusion detection, Cloud, Safety of image processing data, Case study of Fruit images

1. Introduction

As a result of the sharp rise in network attacks over the past few years, researchers' interest in network intrusion detection has risen. Private and governmental organisations rely more and more on their computer networks, thus protecting them from assault is a critical matter. The most important instruments for ensuring safety in computer and network systems are intrusion detection systems. because a single computer network intrusion might result in a significant loss or make the network's consistency unstable. As a result, precise network

* Corresponding author: samata.kancherla@gmail.com

attack detection is crucial. Programmers have been securing their organizations with procedures that recognise and stop definite events for fifty years. Though the development of automated and flexible IDS is urgently required due to the nature of current and emerging threats. As a result, a valuable tool for intrusion detection system applications has arisen in the form of an artificial neural network that was motivated by the nervous system. It offers an intrusion detection system that complies with the ideal requirements and fixes the problems with traditional IDSs. Investigations into the utilisation of ANN for intrusion detection are still ongoing. We briefly introduce IDSs, artificial neural networks, and ANN approaches to intrusion detection in the sections that follow. Additionally, information on the NN, dataset, system implementation, and testing parameters is offered along with research, development, and implementation. Finally, a summary of this field of study is given, along with suggestions for additional research fields. An intruder is a malicious user who has access to network resources and can create havoc. An IDS is employed to identify unauthorised computer system use or network. There are several ways to respond to a network intrusion, but each one requires an in-depth account of the attack. An intrusion detection model was first put forth by Dr. Dorothy Denning in 1987, and it quickly became well-liked in this field of study. The concept she proposed is essentially the basis of the most widely used intrusion detection systems today. There are three types of intrusion detection systems: host-based, network-based, and vulnerability assessment-based. Pattern recognition, network monitoring, integrated anomaly/abuse detection, misuse detection, and anomaly detection are a few of the methods utilised to satisfy the requirements of an intrusion detection system

1.1 ANN:

In general, artificial neural networks are used to forecast the association between unknown and undefinable data correlations. There are both supervised and unsupervised approaches. Intrusion detection is used to find some stochastic processes that have been misapplied in computer use. For best results, Transaction Suspicion must be ranked consistently. Artificial neural networks can learn on their own given the right pre-training. Many methods, including unsupervised, supervised, and reinforcement learning, are employed for this goal. In supervised learning, the instructor must provide the data with known answers. Unsupervised learning involves feeding data for questions for which the solutions are not predetermined. Based on group behaviour, the ANN will determine some relationships. Using the available data sets, hidden patterns are clustered into sets of items that are separated into groups, forming them into an unknown pattern. The strategy of reinforcement learning is founded on observations. Environment observations are used to inform ANN judgments. Negative observations lead to weight adjustments so that the required decision can be made differently the next time. This paper discusses a few ANN applications for image processing using a case study.
For modelling and optimization, neural networks including CNN, KNN, ANN, fuzzy-logic, SVM, and DT are employed.
The typical unknown relations are established when the data sets are grouped using ANN.

2. Related Work:

2.1 ALGORITHMS: Review of works carried on use Artificial Neural Networks for IDS

Warren McCulloch, a neuro-physiologist first used the artificial neuron in 1943. Processing of interconnected elements and converting them into desired outputs are the primary function of ANN. The features of the elements and the weights attached to their relationships among one another determine the outcome of the transformation. The network can change its connections between nodes to produce the required outputs. It has the ability to fix some of the issues that the other current intrusion detection methods have. Alternate methods include artificial neural networks. The network's adaptability would be the first benefit of using a neural network for intrusion detection. Even if the network data is vague or insufficient, a neural network would be able to analyse it. Additionally, the network could perform a non-linear analysis on the data. Since some network attacks may be carried out in concert by several attackers, It's crucial in particular to be able to analyse data from diverse sources non-linearly. Another advantage is how quickly neural networks work.

Ryan et al. conducted one of the first studies on NN-based intrusion detection in 1998. On a ten-user system, they tested and trained an offline NNIDS. They trained their system using backpropagation and a 2-Layer MLP architecture. The data source for testing and exercise was operating system logs in a UNIX environment. False positive and false negative results were the criteria used to assess the system's effectiveness. They used the PlaNet Neural Network simulator to put their concept into action. The table provides a thorough summary of their work.

Cannady made another try in the same location in 1998. He also used backpropagation for training and a 2-Layer MLP architecture for his system. Network packet capture from Real Secure was used as the foundation for training and testing. The MLP network, which has four fully linked layers, received nine network data packet attributes. He determined performance using training and testing data from the RMSE parameter. The table includes a thorough assessment of his work.

Ghosh et al. presented a host-based IDS in 1999, focusing on the creation of programme profiles and their application to distinguish between the behaviour of reliable software and that of malicious software. On the SUN platform, the system was trained and evaluated using information from the Basic Security Module (BSM). The input vectors for the NN were produced using the BSM and a distance metric. The IDS was an MLP with one hidden layer. how many example strings there were that matched how many input nodes there were. In order to record the temporal location of anomalous events, the Lucky Bucket approach is used. The DARPA database was used for performance analysis. Elman Networks were also employed by Ghosh et al. in 1999, and the outcomes of their research are displayed in the following table.

Self-organizing neural networks were investigated in a different study by Rhodes et al. released in 2000 to take advantage of anomalies in network data streams. In contrast to earlier methods that processed the extent of a computer system or network to discover irregularities using self-organizing maps, the suggested strategy deconstructs the system using a succession of more specialised maps. Each neural network in a monitor stack was trained to become somewhat of an expert in recognising usual protocol behaviour and sounding an alarm when that profile is violated. During the test, the invasions were attempts at buffer overflows. In the table, a summary of their literary works is shown.

In order to create a misuse detection model utilising neural networks, Lippmann and Cunningham of the MIT Lincoln Laboratory searched the network traffic for words specific to attacks in 2000. To recognise the hazards of server root privilege and Unix host attacks, they used an MLP network. They fed the neural network phrase counts from network traffic that was unique to an assault. The system employed two neural networks, one for categorising attacks and the other for calculating attack likelihood. With k input nodes, 2k hidden nodes, 2 outputs (normal and attack), and backpropagation as the system's training technique, a two-layer perceptron was created. The results of their study are listed below.
In 2001, Lee and Heinbuch used An element of a neural network hierarchy experimental IDS. Each hierarchy's neural network concentrated on different facets of the nominal TCP behaviour. These TCP operations include port usage, connection establishment, and connection termination. The system was taught three different attack types: SYN flood, quick SYN port scan, and stealth SYN port scan. System training was finished using the backpropagation learning technique. Each neural network's input vectors were produced at random.
When Jirapummin et al. introduced an alternative method for visualising intrusions using SOM and discovering intrusions using resilient propagation in 2002, they did so using the KDD Cup 1999 data set. They used three different forms of intrusions: Satan attacks, Portsweep, and Neptune assault (SYN floods) (port scanning). RPROP uses a three-layer NN architecture, first hidden layer contains 70 nodes, the second hidden layer contains 12 neurons, and the output layer contains 4 neurons. Tan-sigmoidal, log-sigmoidal, and log-sigmoidal transfer functions, respectively, are incorporated into RPROP's output layer, first hidden layer, and second hidden layer.
In 2002[27]28], Bivens et al. proposed a neural network model for a network-based intrusion detection system. Their anomaly detection system used the MLP network. Data from the tcpdump has been read by the system. A summary of their work is provided in the table.
BIn 2007 Iftikhar et alwork .'s in the field of intrusion detection was done [5] [26]. They used the Kddcup 99 entire feature data set for their system. Utilizing RBPROP NN, they trained and evaluated the network. Iftikhar Ahmad, et al. presented another piece of work in 2008 [3] that benchmarked different backpropagation algorithms. They built their business using MLP architecture.
The feedback above states that there are three different approaches to acquire data: [25] using real traffic, [26] sanitised traffic, or [27] simulated traffic. However, at the MIT lab in the United States, IDS were mostly assessed on the common dataset KddCup99. Various

neural network designs were used to build the intrusion detection systems suggested by various researchers. The major testing parameters were rate of detection, ROC, false positive and false negative findings, and findings.

3. Case Study With Image Processing Data Over Cloud With OEM

Equipment : Color and size grading based on Hunter scale color
Software : Ellipse Truesort$^{TM.}$
Data access: Cloud through remote monitoring support

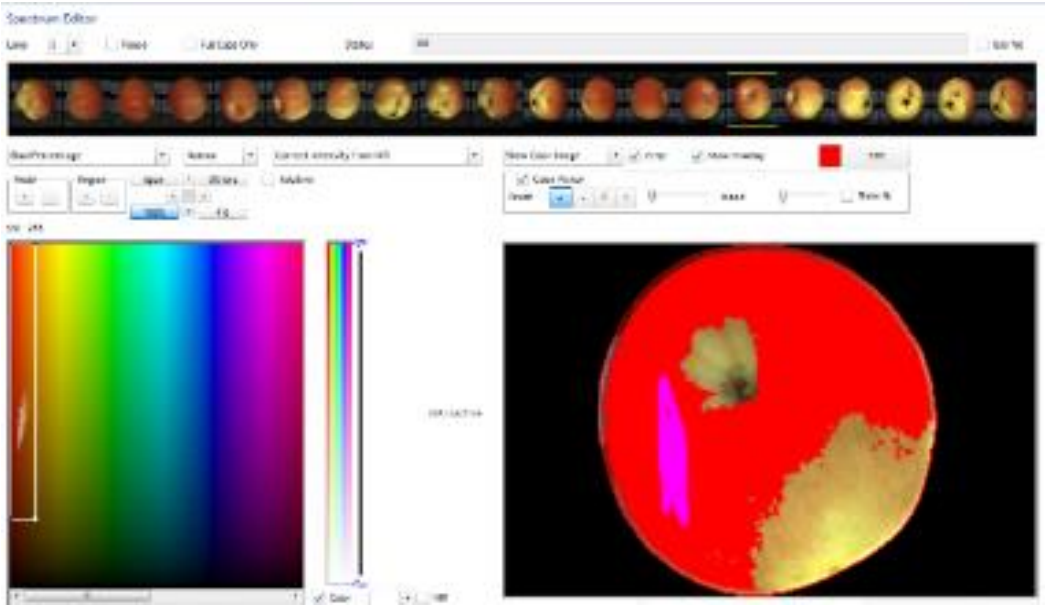

Fig.1. software

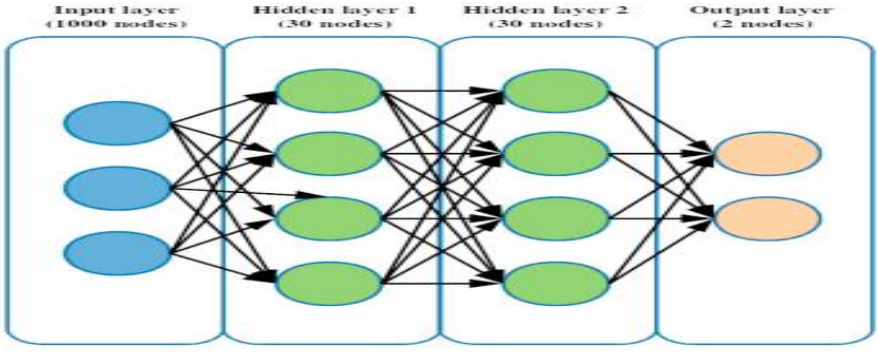

Fig.2. Multilayer Perceptron

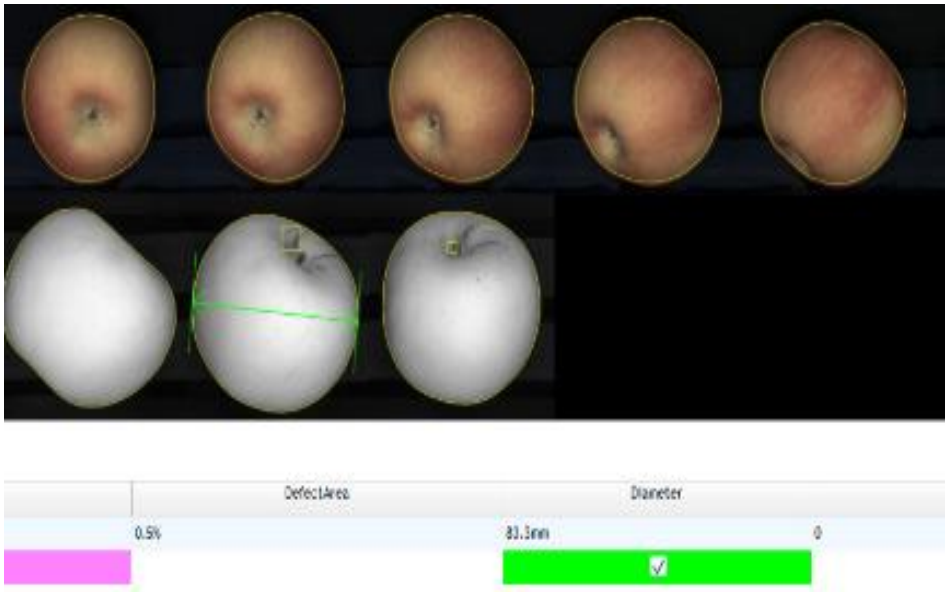

Fig.3. measurement

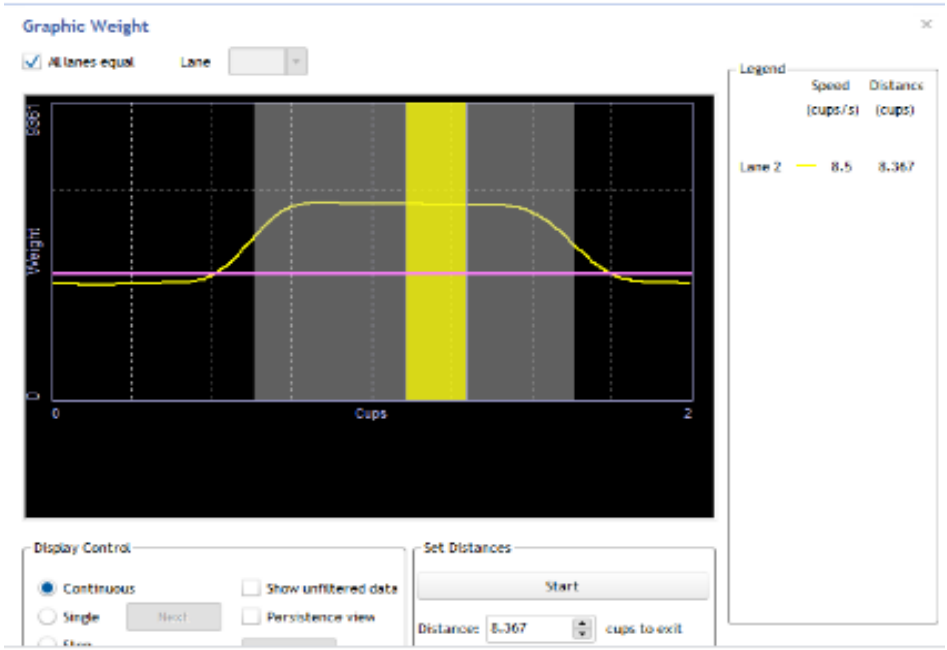

Fig.4.

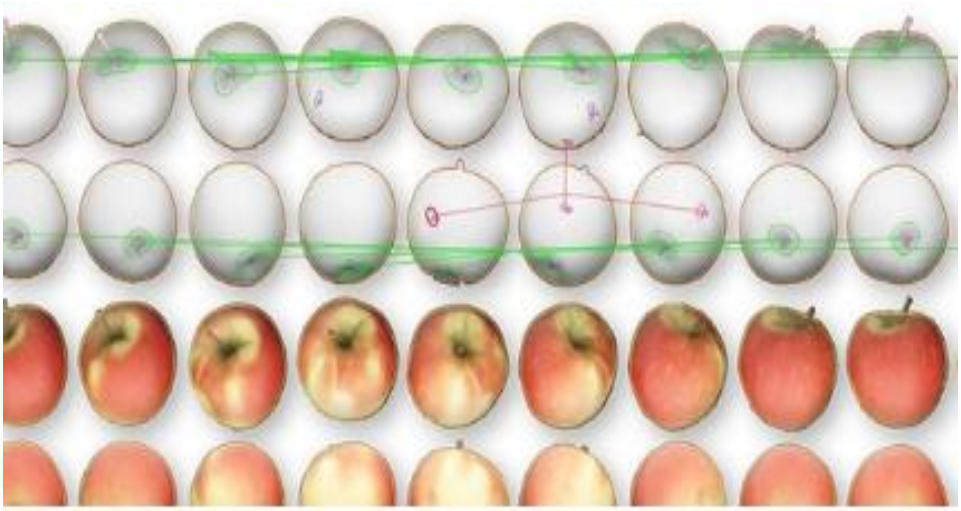

Fig.5.

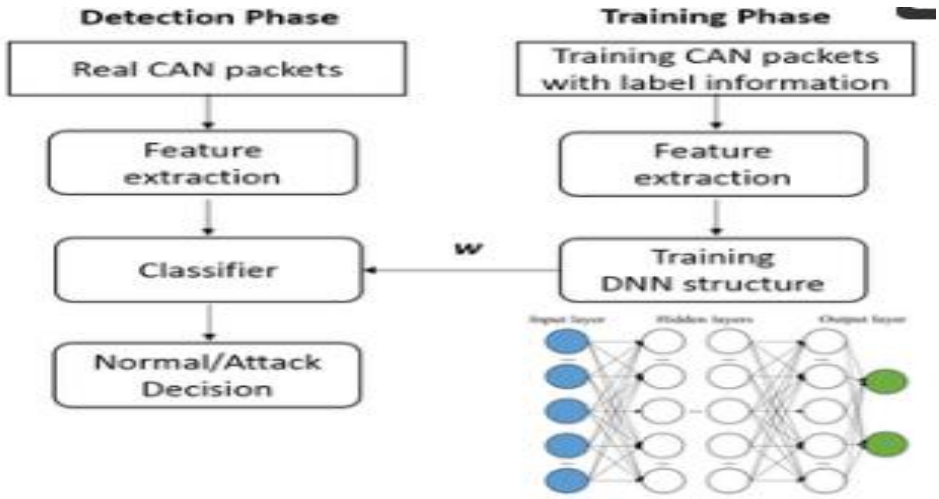

Fig.6.

Some of the deep learning can be incorporated with following neural networks:

[1] Artificial neural networks
[2] Recurrent neural networks
[3] Convolutional neural networks

Several food processing operations can be successfully modeled using ANN. Some of them are discussed below

Sun drying of agriculture produce, for example potato has been modeled using ANN prediction by Tripathy and Kumar. Different number of neurons are used for and optimized the process using four neurons. They used logsig-transfer function and trainrp-

backpropagation algorithm. Error minima calculations were used by them to predict the modeling drying parameters like drying rate and moisture content.

ANN was applied by Icier et al., for modeling the energy parameters during FBD-Fluidized bed drying of potato slices. Several layers of ANN were applied for exergy and energy correlations with respect to the outputs and inputs Drying temperature, time, size of carrot cubes of food and depth of bed. Mean square errors were lowest for the configuration of 28 neurons in hidden layer (Nazghelichi et al.). Total epochs were 2877 for training. Total three times training was given. Coefficient of momentum was 0.66 and size of the step was 0.35.R square values of 0.97 were achieved between energy and exergy prediction.

Momenzadeh et al. used ANN for corn drying process using FBD and microware combination heating. The temperature range were 3 to 60 deg C. with range of wattage of 180 to 900. 170 number of neurons were used for prediction using network with signoidal tangent with hyperbolic function-Tansig and trainrp.

Khazaei et al. reported the modeling of grape dehydration process to predict the water content. Three different layers were required in their model for moisture prediction. Four input nodes were used. The real time prediction was possible with this process.

Aktaş et al. reported the heat pump dehydration process using ANN with back propagation learning algorithm- Fermi transfer function and Levenberg–Marquardt function. High R square were achieved in the range of 0.996 to 0.998.

Mugwort leaves were dried using convection dehydration with air temperature, velocity and time as input and moisture, energy, exergy efficiency and rate of drying as output parameters Karimi et al.

4. Conclusions & Future Scope

Use of ANN in food processing are found to be successful. When the data is shared over cloud, the possibility of hacking can be avoided by devising proper intrusion detection systems. Intrusion primarily helps in identifying what kind of attack is happening on the data can be precisely grouped into clusters and based on identification methods devised, clustering algorithms help in proper detection. This will help in controlling the data theft or mis use. Primary areas of data modeling in food processing found to be frying, drying and dehydration, fluidized bed drying, solar drying, food quality control, color measurement and grading and sorting. Freezing, cooling rates, mass transfer etc., Data intrusion techniques can be easily connected to the already developed ANN based modeled data the compatibility of modeled data of food processing with ANN based intrusion detection will be good. The future work can be taken up to quantitatively correlate ANN based food process modeling data with ANN based intrusion detection systems.

References

[1]Q. Song, Y. J. Zheng, Y. Xue, W. G. Sheng, and M. R. Zhao, "An evolutionary deep neural network for predicting morbidity of gastrointestinal infections by food contamination," *Neurocomputing*, vol. 226, no. Supplement C, pp. 16–22, Feb. 2017.

[2]A. M. Torkashvand, A. Ahmadi, and N. L. Nikravesh, "Prediction of kiwifruit firmness using fruit mineral nutrient concentration by artificial neural network (ANN) and multiple linear regressions (MLR)," *Journal of Integrative Agriculture*, vol. 16, no. 7, pp. 1634–1644, Jul. 2017.

[3]I. P. Gualda, *Aplicação de Redes Neurais Artificiais na Ciência e Tecnologia de Alimentos: Estudo de Casos*, Brazil: University of Londrina, 2008.

[4]K. C. Lai, S. K. Lim, P. C. Teh, and K. H. Yeap, "Modeling electrostatic separation process using Artificial Neural Network (ANN)," *Procedia Computer Science*, vol. 91, no. Supplement C, pp. 372–381, Jan. 2016.

[5]H. Zhang and P. V. Zimba, "Analyzing the effects of estuarine freshwater fluxes on fish abundance using artificial neural network ensembles," *Ecological Modelling*, vol. 359, no. Supplement C, pp. 103–116, Sep. 2017.

[6]A. Giwa, S. Daer, I. Ahmed, P. R. Marpu, and S. W. Hasan, "Experimental investigation and artificial neural networks ANNs modeling of electrically-enhanced membrane bioreactor for wastewater treatment," *Journal of Water Process Engineering*, vol. 11, no. Supplement C, pp. 88–97, Jun. 2016.

[7]E. Disse, *et al.*, "An artificial neural network to predict resting energy expenditure in obesity," *Clinical Nutrition*, Sep. 2017.

[8]Z. Zeković, O. Bera, S. Đurović, and B. Pavlić, "Supercritical fluid extraction of coriander seeds: Kinetics modelling and ANN optimization," *The Journal of Supercritical Fluids*, vol. 125, no. Supplement C, pp. 88–95, Jul. 2017.

[9]B. Dębska and B. Guzowska-Świder, "Application of artificial neural network in food classification," *Analytica Chimica Acta*, vol. 705, no. 1, pp. 283–291, Oct. 2011.

[10]P. Joshi, *Artificial Intelligence with Python*, Birmingham; Mumbai: Packt Publishing, 2017.

[11]K. Warwick, *Artificial Intelligence: The Basics*, 1 edition, London: Routledge, 2012.

[12]R. Milo and R. Phylips, *Cell Biology by the Numbers*, USA: Garland Science, 2015.

[13]C. Wang and W. Slikker, *Neural Cell Biology*, USA: CRC Press, Taylor & Francis Group, 2017.

[14]K. S. Matlin, J. Maienschein, and M. D. Laubichler, *Visions of Cell Biology*, Chicago, USA: University Chicago Press, 2017.

[15]G. Gurkaynak, I. Yilmaz, and G. Haksever, "Stifling artificial intelligence: Human perils," *Computer Law & Security Review*, vol. 32, no. 5, pp. 749–758, Oct. 2016.

[16]A. Amirov, O. Gerget, D. Devjatyh, and A. Gazaliev, "Medical data processing system based on neural network and genetic algorithm," *Procedia - Social and Behavioral Sciences*, vol. 131, no. Supplement C, pp. 149–155, May 2014.

[17]K. Gurney, *An Introduction to Neural Networks*, USA: CRC Press, Taylor & Francis Group, 1997.

[18]J. Waters and T. Wittman, *Quantitative Imaging in Cell Biology*, 1st Ed., vol. 123. USA: Academic Press, 2014
.
[19]H. Esfandian, A. Samadi-Maybodi, M. Parvini, and B. Khoshandam, "Development of a novel method for the removal of diazinon pesticide from aqueous solution and modeling by artificial neural networks (ANN)," *J. Ind. Eng. Chem.*, vol. 35, pp. 295–308, 2016.

[20]P. P. Tripathy and S. Kumar, "Neural network approach for food temperature prediction during solar drying," *International Journal of Thermal Sciences*, vol. 48, no. 7, pp. 1452–1459, Jul. 2009.
[21]M. Azadbakht, H. Aghili, A. Ziaratban, and M. V. Torshizi, "Application of artificial neural network method to exergy and energy analyses of fluidized bed dryer for potato cubes," *Energy*, vol. 120, no. Supplement C, pp. 947–958, Feb. 2017.

[22]T. Nazghelichi, M. Aghbashlo, and M. H. Kianmehr, "Optimization of an artificial neural network topology using coupled response surface methodology and genetic algorithm for fluidized bed drying," *Computers and Electronics in Agriculture*, vol. 75, no. 1, pp. 84–91, Jan. 2011.

[23] P. Helman, G. Liepins, W. Richards, Foundations of intrusion detection (computer security), Year: 1992, Volume: 1, Pages: 114-120. DOI Bookmark:10.1109/CSFW.1992.236783.

[24]Fruit and vegetable grading. https://ellips.com/

[25]Tejasree Ganji, Muni Sekhar Velpuru, Raman Dugyala "Multi variant handwritten telugu character recognition using transfer learning" IOP Conference Series: Materials Science and Engineering Volume:1042 Issue 1 IOP Publishing

[26]Raman Dugyala, N Hanuman Reddy, Shrawan Kumar "Implementation of SCADA Through Cloud Based IoT Devices-Initial Design Steps" 2019 Fifth International Conference on Image Information Processing (ICIIP) pages: 367-372

[27]Nikhitha Pulmamidi, Rajanikanth Aluvalu, V Uma Maheswari "Intelligent travel route suggestion system based on pattern of travel and difficulties" IOP Conference Series: Materials Science and Engineering volume :1042, IOP Publishing

[28]Krishna Keerthi Chennam, Rajanikanth Aluvalu, V Uma Maheswari "Data Encryption on Cloud Database Using Quantum Computing for Key Distribution" Machine Learning and Information Processing, Pages:309-317 , Springer, Singapore