Phoenix: International Multidisciplinary Research Journal
Vol 3, No. 3, July-Sep, 2025
ISSN: 2583-6897

FACE RECOGNITION WITH BILINEAR CNNS: A REVIEW

Dikshendra Daulat Sarpate
Professor, Department of Artificial Intelligence & Data Science, ZEAL College of Engineering & Research, Pune, Email id -dikshendra@gmail.com
B. Sankaraiah
Assistant Professor, Department of Computer Science & Engineering, Malla Reddy (MR) deemed to be University, Hyderabad. Email id.: Shankar61186@gmail.com
B.Rani
Assistant Professor, Department of Computer Science & Engineering, Malla Reddy (MR) deemed to be University, Hyderabad. Email id.: rani@mrec.ac.in
Vemula Nikitha
Assistant Professor, Department of Computer Science & Engineering, Malla Reddy (MR) deemed to be University, Hyderabad. Email id.: nikitha479@gmail.com
Dr. Syed Umar
Professor, Department of Computer Science & Engineering, Malla Reddy (MR) deemed to be University, Hyderabad. Email id.: syedumar@mrec.ac.in

Abstract— The identification of faces from real data, captured images, sensor photos, and database images is a challenging endeavor. This is because there is a wide range of facial appearances, the illumination effect, and the intricacy of the image backdrop. The use of image processing and biometric technology for facial recognition is one of the most effective and significant applications of these technologies. This research presents a discussion of face identification methods and algorithms that have been developed by a number of researchers using artificial neural networks (ANN). These algorithms have been utilized in applications that are not related to image processing and pattern recognition. This article will also discuss how artificial neural networks (ANN) will be utilized for facial recognition, as well as how it is superior to other approaches in terms of efficiency. In order to present an overview of face recognition using ANN, there are many different ways that are provided by ANN. As a consequence of this, this research includes an exhaustive analysis of face recognition experiments and systems that make use of a wide variety of artificial neural network (ANN) methodologies and algorithms. Through the course of this study effort, both the positive and negative aspects of the aforementioned literature studies and systems were incorporated, in addition to a performance examination of a number of different ANN techniques and algorithms. During the past twenty years, the discipline of shape detection has emerged as the most exciting sector of the scientific community. Object identification is the focus of our discussion as we provide an innovative approach to identifying shapes that are similar to one another. The process of determining the correspondences between the vertices of two shapes and then using those correspondences to estimate an aligning transform prior to determining the degree of similarity between the shapes. In this article, we will go over some of the most important aspects of face detection, which are useful in a wide range of applications. These applications include face recognition, image classification, the ability to track faces, facial feature extraction, person identification, identification number, record keeping and access control, grouping, biometric scientific knowledge, human computer interaction (HCI) system, electronic beauty products, and many more. Prior to that, I would like to go over some well-known facial recognition techniques and then some image processing approaches. This is because we won't be able to effectively identify someone unless we remove the primary parts of their face, which are their eyes, nose, and mouth. Calculating the difference between the two shapes involves determining the percentage of the matching mistake that exists between locations that are comparable. As a method for classifying recognition, we make use of a closest neighbor classification.

Facial, Recognition Biometrics, Image Processing, Problem Solving, and Artificial Neural Networks are some of the terms used in the index.

I. INTRODUCTION

Within the context of the present digital era, the human race possesses exceptional recognition abilities. Since the beginning of time, humans have had the ability to recognize and distinguish between various face characteristics. Recent research has shown that computers are capable of performing the same tasks. In the middle of the 1960s, researchers started working on the idea of using computers to recognize human faces. Since that time, face recognition has come a long way, and both public and commercial organizations have started to use the technology. In order to enhance the capacities of scientists working in this field, numerous improvements are now being implemented. The reliable detection of human faces in arbitrary contexts is the initial step in the process of identifying people, and many experts consider that this is the most important process involved in facial recognition. In order to attain a low false negative error rate, the more conventional face recognition algorithms help to reach this goal. These systems implicitly assume that the losses of all false negatives are the same with each other. According to Ouyang Z (2018), human figures are nearly equal in terms of global qualities such as size, aspect ratio, and the position of major features [1]. However, the specifics of human figures can vary greatly within individuals, depending on factors such as gender, race, or body language. There is a vulnerability index at the imaginary location in every single scenario. An image of a card counter, thief, or prohibited individual is taken from the video surveillance system at the Montecito Hotel and Casino. The image is then scanned through a database in order to locate a match and identify the individual. In an effort to limit the amount of criminal activity in the area, the Chicago Police Department installed facial recognition technology-equipped police cameras in the City entertainment zone in the year 2022 [2]. However, the program proved to be ineffective, and it was eventually dismantled in the year 2022 [3].

II. FACE RECOGNITION SYSTEM

Facial, identification processing, extraction of features, and face recognition are the four fundamental elements of every biometric system, as depicted in Figure 1.

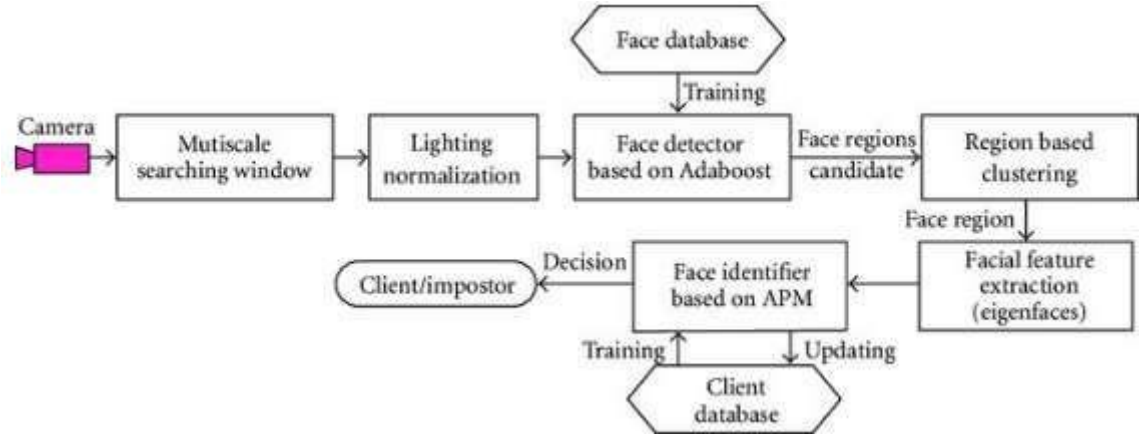

Fig 1: Architecture of Face Recognition System [4]

As demonstrated in Figure 1, the initial responsibility of a biometric is to capture an image through the use of a video camera, a camera, or a computer. This image is then sent to the subsequent stage of the face recognition system, which will be addressed in the following aforementioned section [5,6]:
Face Recognition: The major objective of this stage is to identify a face in an image that has been recorded or in an image that has been selected from the library. Using this face identification approach, the given image is examined to determine whether or not it contains a face image. Once the face has been identified, the output is then sent to the pre-processing stage.

Pre-processing: This stage serves as a face identification pre-processing step, and it is during this stage that undesirable noise, blur, shifting lighting conditions, and shadowing effects are eliminated by the application of pre-processing techniques. Following the acquisition of a facial image that is both fine and smooth, the technique for feature extraction will be carried out.

Face extraction: The use of a feature extraction algorithm is required in order to accomplish the task of face extraction. Through the use of extractions, many tasks such as data packing, dimensionality reduction, salient extraction, and noise cleansing can be accomplished. Following the completion of this phase, a face patch is often transformed into a corrected vector or a collection of feature vectors along with their respective positions.

Face Recognition: The process of facial identification can be broken down into two distinct categories: local feature-based methods and global functionality methods when it comes to facial recognition. The classification of human faces can be done according to both local and global characteristics. As a result of the fact that global traits are less discriminatory than localized qualities, it is simpler to capture them. Local facial characteristics have the potential to be extremely discriminatory; yet, they are susceptible to being damaged by alterations in the facial appearance or partial face blockage. Face recognition systems of today recognize faces by utilizing numerous perspectives of faces; some authors have offered multi-view face recognition algorithms for the purpose of recognizing each facet of a head, including the left, right, front, top, and bottom of the head.

III. Face Detection Technique

A number of different representation strategies are utilized in the process of face identification. These strategies include knowledge-based, features sub-distributors, template matching method, image methods, and part-based methodologies.

There are methods for locating faces that are known as invariant features. These methods involve locating structural features that are present independent of the position, viewpoint, or lighting conditions. In order to differentiate themselves from methods that are based on experience and understanding, feature invariant approaches begin with the extraction of features and the identification of face candidates. They then verify each candidate by using spatial relations among these features. On the other hand, knowledge-based methods typically use the entire image and are sensitive to complex ethnicities and other variables. There are further works that can be found in [7,8]. The detection of faces based on color data and random labelled graph fitting are both included in this description.

The knowledge that humans have [9] about what constitutes a typical face is encoded through the use of knowledge-based techniques. In most cases, the rules are able to capture the links that exist between the characteristics of the face. These algorithms are primarily designed for face localization, which refers to the process of detecting the position of a single face in a photograph.

Various techniques for matching templates In order to characterize the face as a whole or each facial feature on its own, typical patterns of a face are maintained. It is necessary to calculate the associations between being a source image and the stored patterns in order to do detection analysis. Through the utilization of these methods, both face detection and characterization have been successfully performed. It is in this context that deformable pattern recognition [10] comes into play. This technique allows for the templates of faces to be distorted in accordance with specific criteria and constraints.

The appearance-based approaches, as opposed to the pattern matching approaches, learn models (or templates) by employing a sequence of training photos that are intended to capture the representational variability in facial characteristics. The detection process then makes use of the learnt models after that has been completed. There are further noteworthy methods that can be found in [11]. The utilization of view-based facial recognition is an illustration of this method.

As a result of the development of the visual theoretical framework and point of interest detectors, such as the differential of Gaussian sensor [7] (which is utilized in the SIFT detectors) and the Gaussian curved protection system, part-based approaches have recently garnered a great deal of attention. Both the identification of faces via the use of a generating theoretical framework and the recognition of elements of faces through the use of an SVM classifier are well-known methodologies in this category.

IV. NEURAL NETWORK TECHNIQUE

A neural network is a categorizing technique that is both very effective and dependable. It may be utilized to make predictions about unknown data in addition to data that is already known. In addition to being effective with linear and nonlinear datasets, it also works well with separate datasets. Neural networks have been put to use in a wide range of applications, some of which include the interpretation of visual scenes, the processing of natural language, the identification of faces, the detection of fingerprints, and the recognition of iris patterns, amongst others [12].

Artificial neural networks, also known as ANNs, are constructed from a network of biological neurons that are referred to as "modules." A number is assigned to each of these networks based on the strength of their interconnections, which can be either inhibition (with a maximum of -1.0) or stimulation (with a maximum of +1.0). Together, these networks are connected to one another.

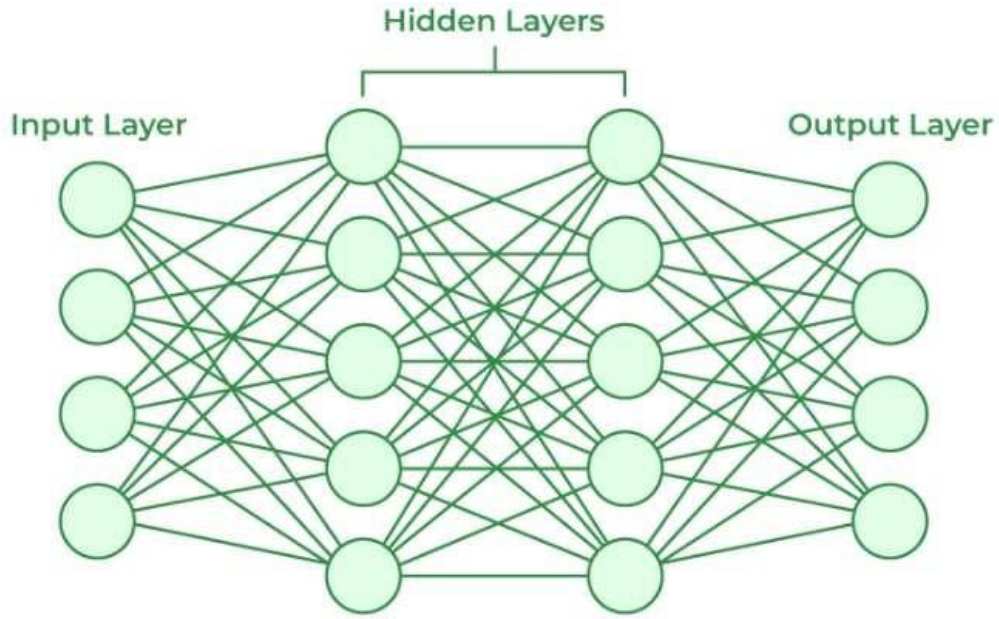

Fig 2: Artificial Neural Network

Single Layer Feed Forward Network: A feed technique or acyclic net is a genetic algorithm in which the input layer of destination node is linked to an outputs input neurons but not the other way around. As indicated in Figure 3, "single layer" refers to the output layer of compute nodes in a monolayer net.

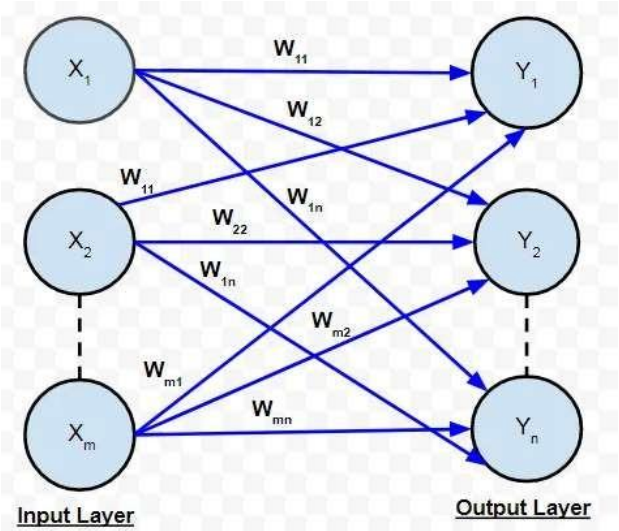

Fig 3: Single Layer Feed Forward Network

Multilayer Feed Forward Network: Hidden neurons or hidden units are used as computing nodes in this type of network, which comprises one or more hidden layers for its computations. It is the hidden neurons that are accountable for the interaction that takes place between the network and the external input. In the network's input layer, the destination node is responsible for delivering the input signal to the neurons that are located in the second layer of the first hidden layer. The features that were retrieved from the second layer are being used as input for the generation of the next layer, and so on... The entire response of the network to the face detection that was provided by the destination node in the first layer of input is made up of the number of output signals that were produced by the neurons in each layer [13].

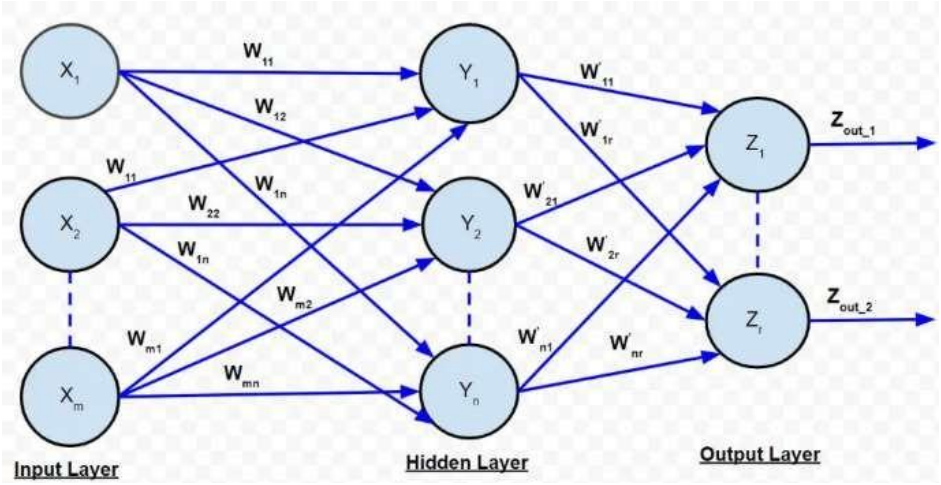

Fig 4: Multilayer Feed Forward Network

Recurrent Network: An example of a feed forward neural network that contains one or more hidden layers and at least one feedback mechanism is shown in Figure 5. This network is a function of the length. One possible kind of return is self-feedback, which occurs when the output of a neuron is delivered back into the cell's own input. With the condition that the neural network contains negative units, dynamical behavior is the result of the utilization of unit delay elements in feedback loops. Other types of networks include the delta-bar-delta, Hopfield, learning vector, counters propagating, probabilistic, Hamming, Boltzman, unidirectional auditory processing, spacio-temporal patterning, adaptive resonant, self-organizing map, and recirculation networks, amongst others [14].

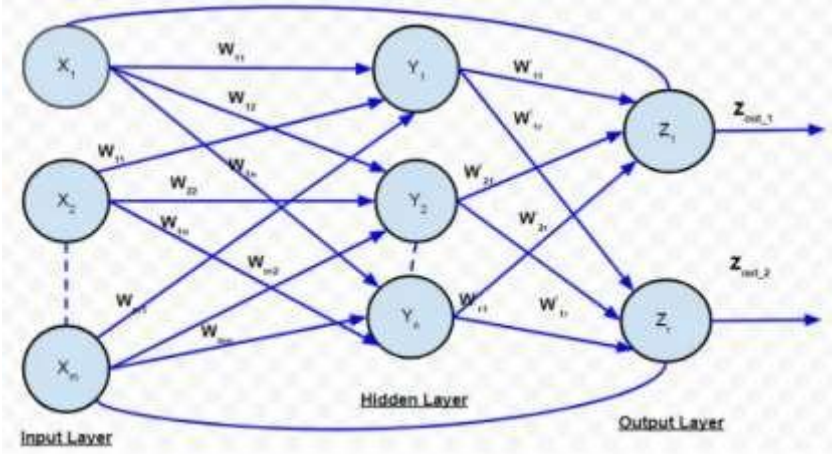

Fig 5: Recurrent Connected Network

ANNs can be broken down into three distinct categories: single layer feed forward neural networks, layered feed forward neural networks, and recurrent neural networks. Each of these categories is going to be reviewed in this article. In the process of face recognition, numerous researchers have utilized these kinds of artificial neural networks (ANN), and they have offered a variety of face recognition algorithms.

V. CONVOLUTIONAL NEURAL NETWORK (CNNS)

Each convolutional, pooling, and non-linear layer is a component of a convolutional neural network (CNN), which is a type of neural network that is composed of a sequence of units. Deep convolutional neural networks (CNNs), which typically consist of ten or more of these units and are trained on massive labelled datasets such as ImageNet, have recently produced generic features that can be utilized in a variety of recognition tasks. These tasks include image processing [15], object tracking [6], feature extraction [8], and texture acknowledgement [3]. These structures, when combined with detectors that pinpoint specific portions of the item, have also delivered state-of-the-art results in the arena of fine-grained recognition. For example, they have been able to recognize the breed of a dog, the species of a bird, or the model of a car. When CNNs are not equipped with component localization, they frequently perform badly. This is due to the fact that there is a significant amount of variation in appearance due to the different poses that instances of a category can assume. The posture variation, which is also a phenomenon seen in the face identification problem, is so prevalent that it completely obscures the minute differences that exist between the categories. These methods, on the other hand, have the drawback of necessitating (a) the human annotation of parts, which can be a hassle at times, and (b) the detection of components, which can be expensive from a technological standpoint.

VI. BILINEAR CNNS

This is Tsung-Yu Lin Subhranshu. It was Aruni RoyChowdhury. It's Maji. It is Johan. The Bilinear CNN strategy for facial recognition was presented in [9], which has shown surprising greater performance on fine-grained recognition issues. Managed to learn [9] introduced this approach. In contrast to CNN models that are based on parts, it bridges the gap between texture models. The outputs of two CNNs' convolution - necessary specification are mixed using the outer products at each place on the image. This process is repeated at each location. The piecewise linear feature that was generated subsequently is applied to the complete image, which ultimately results in an ordered less description. It is possible to normalize this matrix in order to add more invariance [16, 17]. If one of the feature extractions was a part detector and another was computed local features, then the bilinear vector that was produced can be used to characterize the representatives of a part-based description. On the other hand, bilinear vectors are comparable to Fisher vector calculations in the sense that local attributes are combined with soft membership to a collection of cluster centers through the use of outer products.

VII. LONG SHORT-TERM MEMORY (LSTM) NETWORKS

LSTM networks, also known as Long Short-Term Memory networks, are a specific kind of recurrent neural network (RNN) that was developed with the intention of identifying long-term dependencies in sequential data patterns. LSTMs, in contrast to standard feedforward networks, are equipped with memory cells and gates that give them the ability to selectively remember or forget information throughout the course of processing time [17]. Speech recognition, natural language processing, time series analysis, and translation are just some of the tasks that LSTMs are particularly effective at performing because of this capabilities.

By utilizing methods such as the forget gate, input gate, and output gate to limit the flow of information, LSTM networks are able to handle the issues that typical RNNs face, such as vanishing or bursting gradients, with exceptional success. On the other hand, choosing the suitable architecture and parameters for LSTM networks continues to be a difficult task. These networks need to be meticulously configured by researchers and practitioners in order to get the best possible performance.

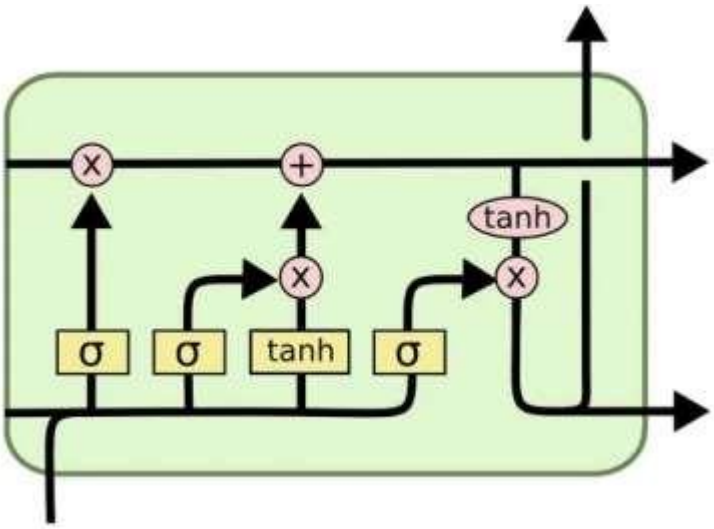

Fig 6: LSTM Network

VIII. Conclusion

This paper includes a review of the existing research on facial recognition systems that are based on artificial neural networks (ANNs). In this research, several architectures, methodologies, algorithms, methods, databases for training or testing images, and performance measurements of face recognition systems were discussed for each of the studies conducted. Every researcher has their own methodology for detecting faces from databases or films, and many studies have sought to tackle the problems related with the previously stated method, however these methods still have certain advantages and downsides.

.

References

[1] Ouyang Z., Cui G., Zhao J., and Liu Q., Image depth-of-field rendering algorithm based on hierarchical anisotropic filtering, Optics Technology. (2018) 44, no. 4, 469–475.

[2] Naveen Sai Bommina, Uppu Lokesh, Nandipati Sai Akash, Dr. Hussain Syed, Dr. Syed Umar, "Optimized AI Models for Real-Time Cyberattack Detection in Smart Homes and Cities", International Journal of Applied Engineering & Technology, Vol. 4 No.1, June, 2022.

[3] Habeeb, M. S., & Babu, T. R. (2022). Network intrusion detection system: a survey on artificial intelligence-based techniques. Expert Systems, 39(9), e13066.

[4] Liu H., Hu X., and Xu L., Research on remote real-time rendering system based on graphics cluster, Journal of System Simulation. (2019) 31, no. 5, 886–892.

[5] Umar, Syed, Bommina Naveen Sai, Nagineni Sai Lasya,Doppalapudi Asutosh, and LohithaRani. "Machine Learning based Sentiment Analysis of Product Reviews Using DeepEmbedding." Journal of Optoelectronics Laser 41, no. 6(2022): 108-113.

[6] Habeeb, M. S., & Babu, T. R. (2024, October). Enhancing IoT Security Through Advanced Feature Selection and Deep Learning. In International Conference on Computing and Communication Networks (pp. 37-49). Singapore: Springer Nature Singapore.

[7] Chen Y.-C., Patel V. M., Phillips P. J., and Chellappa R., Dictionary-based face and person recognition from unconstrained video, Ieee Access. (2015) 3, 1783–1798, https://doi.org/10.1109/access.2015.2485400, 2-s2.0-84959857666.

[8] Nandipati Sai Akash, Naveen Sai Bommina, Uppu Lokesh, Hussain Syed, Syed Umar, "Optimized Block Chain-Enabled Security Mechanism for IoT Using Ant Colony Optimization", International Journal on Recent and Innovation Trends in Computing and Communication, (2023), 11(10), 1226–1233.

[9] K Sankar, Divya Rohatgi, S Balakrishna Reddy, "COX Regressive Winsorized Correlated Convolutional Deep Belief Boltzmann Network for Covid-19 Prediction with Big Data", Grenze

International Journal of Engineering & Technology (GIJET), Grenze ID: 01.GIJET.9.1.547, © Grenze Scientific Society, 2023.

[10] Naveen Sai Bommina , Nandipati Sai Akash, Uppu Lokesh , Dr. Hussain Syed , Dr. Syed Umar, "Multi-Objective Genetic Algorithms for Secure Routing and Data Privacy in IoT Networks", International Journal of Communication Networks and Information Security (IJCNIS), (2020), 12(3), 632–643.

[11] RS Supriya Khaitan, Divya Rohatgi, Sana Nalband, Tejali Mhatre, Shweta Patil, "Enhancing Essay Grading Efficiency and Consistency through Two-Layer LSTM Models and Attention Mechanisms", Journal of Information Systems Engineering and Management 10 (2), 191-202.

[12] Ahmad, Z., Khan, A. S., Aqeel, S., Julaihi, A. A., Tarmizi, S., Annuar, N., & Habeeb, M. S. (2022, May). S-ADS: spectrogram image-based anomaly detection system for IoT networks. In 2022 Applied Informatics International Conference (AiIC) (pp. 105-110). IEEE.

[13] Naveen Sai Bommina, Nandipati Sai Akash, Uppu Lokesh, Dr. Hussain Syed, Dr. Syed Umar, "A Hybrid Optimization Framework for Enhancing IoT Security via AI-based Anomaly Detection", International Journal on Recent and Innovation Trends in Computing and Communication, (2023) ISSN: 2321-8169 Volume: 11 Issue: 3.

[14] M. Mukhedkar, D. Rohatgi, V.A. Vuyyuru, K.V.S.S. Ramakrishna, Y.A. Baker El-Ebiary, V.A. Asir Daniel, "Feline wolf net: A hybrid lion-grey wolf optimization deep learning model for ovarian cancer detection", Int. J. Adv. Comput. Sci. Appl., 14 (9) (2023)

[15] Uppu Lokesh , Naveen Sai Bommina , Nandipati Sai Akash , Dr. Hussain Syed , Dr. Syed Umar. (2021). Deep Reinforcement Learning with Genetic Algorithm Tuning for Intrusion Detection in IoT Systems. International Journal of Communication Networks and Information Security (IJCNIS), 13(3), 582–595. [10] P. Grother and M. Ngan. Face recognition vendor test (FRVT): Performance of face identification algorithms. In NIST Interagency report 8009, 2014.

[16] D Veerendra, BN Umesh, A Khandare, D Rohatgi, K Tiwari, S Datta, "ECA-MURE algorithm and CRB analysis for high-precision DOA estimation in coprime sensor arrays", IEEE Sensors Letters 7 (12), 1-4.

[17] Uppu Lokesh, Naveen Sai Bommina, Nandipati Sai Akash, Dr. Hussain Syed, Dr. Syed Umar, "Designing Energy-Efficient and Secure IoT Architectures Using Evolutionary Optimization Algorithms", International Journal of Applied Engineering & Technology, Vol. 4 No.2, September, 2022.