IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 7, 2019, accepted September 22, 2019, date of publication October 4, 2019, date of current version November 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2945545

Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition

OLUDARE ISAAC ABIODUN[1,2], AMAN JANTAN[1], ABIODUN ESTHER OMOLARA[3], KEMI VICTORIA DADA[4], ABUBAKAR MALAH UMAR[5], OKAFOR UCHENWA LINUS[6], HUMAIRA ARSHAD[7], ABDULLAHI AMINU KAZAURE[8], USMAN GANA[9], AND MUHAMMAD UBALE KIRU[10]

[1]School of Computer Sciences, Universiti Sains Malaysia, Minden 11800, Malaysia
[2]Bingham University, New Karu, Nigeria
[3]Department of Computer Science, Ahmadu Bello University, Zaria 810107, Nigeria
[4]Department of Statistics, Ahmadu Bello University, Zaria 810107, Nigeria
[5]Department of Nuclear Science and Technology, Energy Commission of Nigeria, Abuja 900211, Nigeria
[6]Department of Mathematics, Nigerian Defence Academy, Kaduna 800213, Nigeria
[7]Faculty of Science, The Islamia University of Bahawalpur Pakistan, Bahawalpur 63100, Pakistan
[8]Hussaini Adamu Federal Polytechnic, Kazaure 733, Nigeria
[9]Department of Computer Science, Ibrahim Badamasi Babangida University, Lapai 911101, Nigeria
[10]School of Computer Science and Information Technology, Lovely Professional University, Phagwara 144411, India

Correspondence authors: Aman Jantan (aman@usm.my) and Oludare Isaac Abiodun (aioludare@gmail.com)

This work was supported in part by the Universiti Sains Malaysia Research University (RUI), Malaysia, under Grant 1001/ PKOMP/ 8014017, and in part by The Center for Cyber Safety and Education, United States Internal Revenue Segregated Fund of $(ISC)^2$, Inc. Code. EIN: 45-2405127 through the $(ISC)^2$ Graduate Cybersecurity Scholarship Award, Clearwater, FL, USA.

ABSTRACT The era of artificial neural network (ANN) began with a simplified application in many fields and remarkable success in pattern recognition (PR) even in manufacturing industries. Although significant progress achieved and surveyed in addressing ANN application to PR challenges, nevertheless, some problems are yet to be resolved like whimsical orientation (the unknown path that cannot be accurately calculated due to its directional position). Other problem includes; object classification, location, scaling, neurons behavior analysis in hidden layers, rule, and template matching. Also, the lack of extant literature on the issues associated with ANN application to PR seems to slow down research focus and progress in the field. Hence, there is a need for state-of-the-art in neural networks application to PR to urgently address the above-highlights problems for more successes. The study furnishes readers with a clearer understanding of the current, and new trend in ANN models that effectively addresses PR challenges to enable research focus and topics. Similarly, the comprehensive review reveals the diverse areas of the success of ANN models and their application to PR. In evaluating the performance of ANN models, some statistical indicators for measuring the performance of the ANN model in many studies were adopted. Such as the use of mean absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE), and variance of absolute percentage error (VAPE). The result shows that the current ANN models such as GAN, SAE, DBN, RBM, RNN, RBFN, PNN, CNN, SLP, MLP, MLNN, Reservoir computing, and Transformer models are performing excellently in their application to PR tasks. Therefore, the study recommends the research focus on current models and the development of new models concurrently for more successes in the field.

INDEX TERMS Artificial neural networks, application to pattern recognition, feedforward neural networks, feedback neural networks, hybrid models.

I. INTRODUCTION

Artificial neural networks (ANNs) are referred to as non-linear statistical data models that replicate the role of biological NNs [1]. Statistical pattern approach has been the most commonly studied and utilizes in practice [2]. However, an artificial neural network (ANN) models have been attractive [3], [4]. ANNs are increasingly attractive, effective, efficient, and successful in achieving pattern recognition (PR) in many problems [5], [6].

Unlike conventional pattern approaches, ANN can easily model complex or multi complexes task [7], [8]. The former

The associate editor coordinating the review of this manuscript and approving it for publication was Rajeeb Dey.

conventional techniques applied to handle PR problems are classified into structural, statistical, and hybrid approaches [9]. However, both the statistical and structural approaches can produce unsatisfactory results if they are applied as a solution to complex PR problems only. For instance, in the application, the structural method can be weak and not be able to perform in handling noise patterns. Similarly, it can be weak and ineffective in resolving numerical semantic information challenges.

Likewise, the statistical method is incapable of using information concerning patterns structures. Thus, the intuiting of both approaches combined attracted research attention which gives rise to a hybrid approach. However, nowadays the ANN models are used because they can yield a better result in PR problems even in multi complexes tasks.

The function of ANN in PR is unique and flexible with remarkable success. PR is a computational paradigm used for the classification of raw data. PR embrace a plethora of approaches that provide the development of different applications in various field of endeavor. The practicability of these approaches is the intelligent human imitation.

A pattern can be referred to as a set of items, objects, images, events, cases, situations, features or abstractions where facets of a set are alike in an unequivocal sense. According to Norbert Wiener, ''Pattern is an arrangement, it is characterized by the sequence of the features of which it is made-off instead of inherent underlying of features,'' [10]. Whereas, Watanabe defined a pattern as ''an entity'' [10].

It can also be defined by the unique or recurrent denominator amidst multiple samples of an entity. For example, common things in fingerprint images can define a pattern of a fingerprint. Hence, a pattern can either be a fingerprint image, a human face, a handwritten joined word, a barcode, Internet web page, or a speech signal while recognition is a task of identifying an object, feature, or event.

Thus, this paper identified the current problems of ANN in PR. Likewise, it presents groundbreaking results of studies from the comprehensive analysis of ANN application to PR. Moreover, it highlighted many author's opinions of ANN's application to PR. It acknowledges the foundational development of a self-activating PR system. It addresses advanced applications and innovative in recognition using artificial neural networks.

Finally, it discusses future technological perspectives of ANN's application to PR. No doubt that many kinds of research are sponsored to harness the potentials in artificial intelligence (AI), which brought about substantial growth in recent times.

II. ANN PATTERN RECOGNITION TASK

In the task of PR, ANN uses knowledge of how man's brain systems execute information. NNs are suited explicitly for pattern interrelation. ANN, in its functionality, provides a paradigm for PR achievement that requires large associated networks of nonlinear and straightforward units know as neural nets. PR task is achieved by applying a feedforward neural network (FFNN) that trained accordingly [11], [12]. An ANN design for PR and signal direction can be represented in Fig. 1.

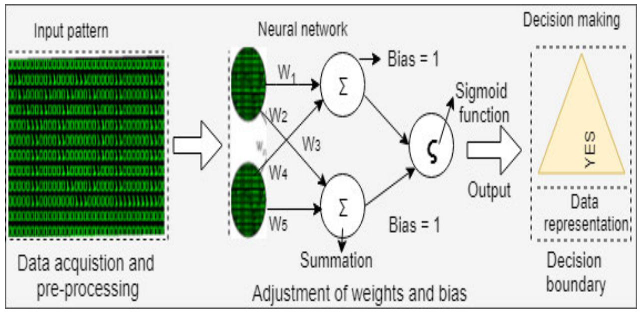


FIGURE 1. ANN model information flow for PR.

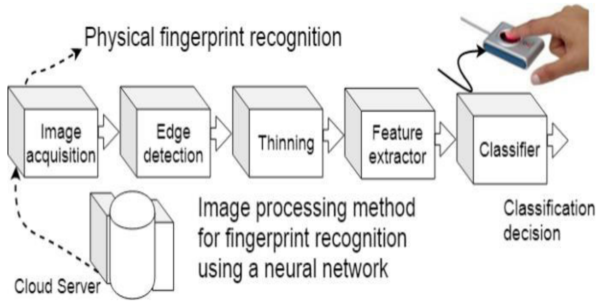


FIGURE 2. Image processing method for fingerprint recognition.

During training, an ANN model such as FFNN trained link input to output design. Potential of FFNs become amazing if it discloses pattern in output unlink to input. Therefore, network proffer output that connects to a trained input pattern minimal from a given pattern. A novel example of an image processing method for fingerprint recognition and biometric identification is demonstrated in Fig. 2.

During the operation, the finger as an image is acquired then converted into an array of numbers which can be controlled by computer.

Edge recognition and thinning are part of the pre-processing steps in removing noise, enhancing the image, when necessary, segmenting the image into parts.

A typical ANN model for the PR task is the use of a convolutional neural network (CNN) or ConvNet. CNN is a group of deep feedforward artificial neural networks (DFFANNs), normally applied to analyze visual imagery [13], [14]. CNN's possesses three structural understanding that brings shift, invariance and distortion; weight replication, and temporal subsampling.

In a finger-vein biometric identification, unlike some existing biometric recognition techniques such as face and fingerprint, the vein patterns are inside the body, making them almost impossible to reproduce. Finger-vein biometrics can be more secure alternative techniques to conventional finger-vein recognition methods. Since finger-vein biometrics are secure without being vulnerable to either forgery, change,

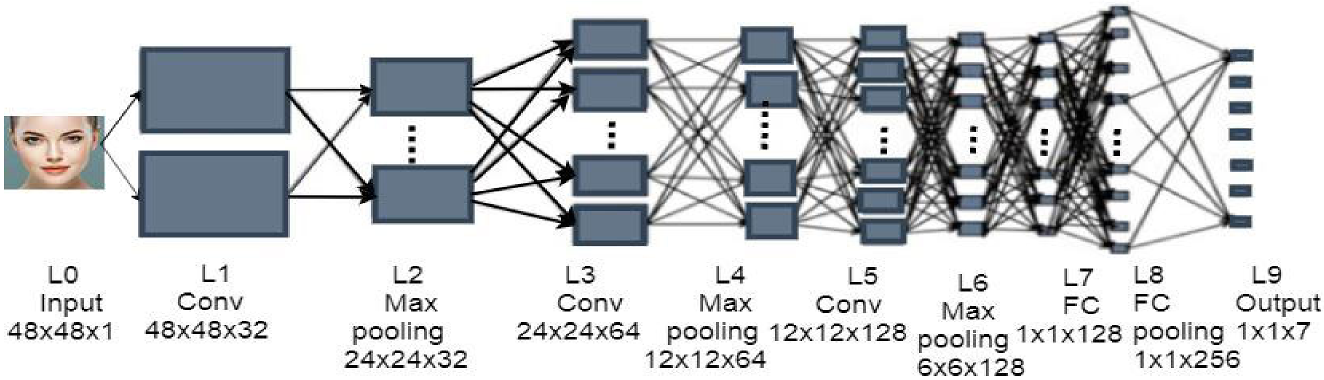


FIGURE 3. A demonstration of pattern recognition in facial expression using CNN's, where L's represent layers.

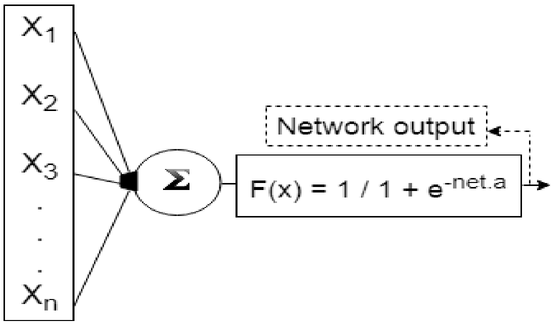


FIGURE 4. Neuron model of pattern recognition.

damage overtime. In conventional finger-vein recognition approaches, complex image processing is needed to eliminate noise and extract and then enhance the features so that the image classification can be carried out to achieved high-level performance accuracy.

Nowadays, CNN has witnessed remarkable success in PR; for example, in an experimental realization of PR with CNN. CNN design applicable to the facial expression [15]–[17] or emotion recognition [18] can be superb. A more novel method of using a CNN can be illustrated in an image recognition as in Fig. 3.

The model of a network consists of analogy cells like a neuron, which has subunits. Sample of these cells is used in a network such as represented in Fig. 4.

In a neuron model of PR, the multi-layer hierarchal network is consisting of a plethora of cell layers [19]. In the multi-layer hierarchal network, there are forward and backward connections between cells. Thus, the network can be trained for the best-matched solution to a given problem.

III. AREAS OF ANN APPLICATION TO PR AND PR PROCESSING

ANN applications to PR include the following area; business, communication, automation, biometrics (face, language, voice, fingerprint, gait, iris recognition), smell recognition (e-nose, sensor network), defect detection in chip manufacturing, speech, signal, and credit application. Others include handwritten (digital and letter or word recognition), image, bioinformatics, biotechnology, data mining, military, crime detection, terrorist recognition, credit fraud detection, interpreting DNA sequences, smell recognition, medical diagnosis, fruit/vegetable detection, a forensic investigation [20]–[23]. Likewise, ANN is useful in crops and animals' yield PR, crops, and animals' species PR, etc [24]–[27].

ANN applies to PR, which can offer a solution in diverse problems such as speech classification, handwritten of characters, and medical diagnosis [28]. ANN in the PR process includes data acquisition and remote sensing such as resolution, bandwidth, and measuring of physical variables, etc.

In ANN application to PR, the process involves preprocessing to post-processing. Preprocessing means elimination of disturbances like noise in data and separation of patterns from an item. Then follow by features extraction (FE), which is the act of finding a new representation of the item.

As mentioned earlier classification takes place in PR. So, in ANN classification is a way of applying features to assign a pattern. Thus, learned ANN models could classify items with identified features. Then, post-processing takes place that is, decision making regarding PR. Stratification can help in knowing the group of new data in learning data analysis procedures [29]–[31]. Dataset inputted into a PR system is shared into two sets, that is, the training set and the testing set. The design of PR systems is shown in Fig. 5.

Though, ANN system may learn effectively from the training set. However, the proficiency of the system is a monitor at a practical or operation stage. A typical component of a PR system in ANN is presented in Fig. 6.

Three elements influence the success of the PR techniques; the volume of data, the method applied, designer, and user. The task in PR is building a system with the capability to handle large data. Means of resolving challenges of PR is in the selection of analysis model like preprocessing, scheme

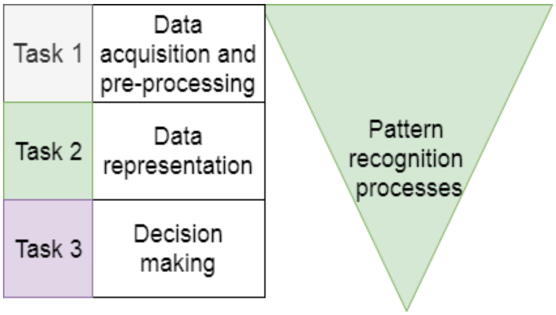


FIGURE 5. Processes of PR.

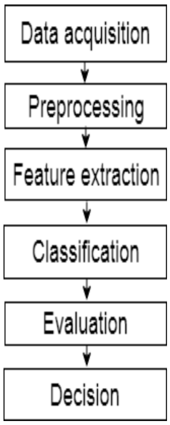


FIGURE 6. Components of ANN PR system.

and post-processing or the decision-making model. More also, learning from a trained set of PR and the desired output of PR is important in ANN [32]–[35].

Interestingly ANN applied to images can improve image restoration like image compression and nosing issues. The work by Wei *et al.* [36]; Dony and Haykin [37] focuses on image improvement through image compression. By their experimental result studies demonstrates the effectiveness of applied ANN models to image compression. Likewise, ANNs class like FFNN useful in diverse areas of remote sensing in agriculture specifically in crop type classification, crop and animals' production estimation. Likewise, agricultural products and data can be predicted through ANN models.

ANN tools provide an easy, simple, and faster method of data analysis in discovering the modern way to success in business, manufacturing, and many realms of life. Also, ANN tools provide across products, channels, and customer achievements.

More also, ANN tools include data cleansing, and it allows opportunities for an organization to tackle management challenges like improving materials, products, services, financial crimes [38] by current means.

ANN is useful in addressing the problems of crime detection and prevention [39] through PR [40]–[44]. With the evolution of connected computers, the Internet, technologies, and digital media [45]. Addition to the deployment of highly entertaining content software [46]–[48]. More success can be achieved without a deep knowledge of the external or internal process of a task [49]–[53] and the detection of anomalous patterns during walking [54].

IV. ANN MODELS PATTERN RECOGNITION PROBLEMS

Despite the long-time existence of ANN, its application to PR still has some problems that are yet to be resolved. These problems include the difficulty of recognizing complex patterns such as object classification, location, scaling, and object with arbitrary orientation, etc., [55]–[57]. Although, there is some attempt in the past to address the challenges in PR tasks using convolutional NN (CNN) and processing of image [58]–[60]. However, more study is requiring in tackling these long-time problems of recognizing complex patterns for more success.

There are two problems associated with the use of ANN in PR. Firstly, problems related to ANN techniques and secondly, specific application problems.

A. CURRENT PROBLEMS RELATED TO ANN TECHNIQUES

Current problems related to ANN techniques are as follow;

1) DATA BEHAVIOR ANALYSIS PROBLEM

Research into the analysis of data behavior such as speed, accuracy, performance, volume, fault tolerance, latency, convergence, and scalability concerning ANN success in PR is yet to be fully understood or studied [61]–[63].

2) SUPERVISED CLASSIFICATION PROBLEM

Supervised ANN is popularly used for addressing classification task, and in some cases, the tasks of classification has been challenging. When there is an input, ANN recognition of pattern entails the task of supervised classification [64]–[67]. The ANN recognition of pattern problem is solely in the classification task [68]–[70], which remains to be addressed successfully. Hence, research is still requiring in the area ANN supervised classification task.

3) SCALING PROBLEM

Normalization or scaling is requiring so that the inputs can be in the comparable range. Scaling is remarkably helpful in the transpose of the input variables to the data range in which the sigmoid activation functions can lie, an example is tanh $[-1, 1]$ and lofistic $[0, 1]$.

Scaling becomes more significant with activation functions sigmoid and tanh. However, both sigmoid and tanh has vanishing gradient problem when applied to activation function during scaling. However, scaling is not a functional requirement for NNs to learn. Therefore, scaling issues with ANN require research attention for a solution because it remains unsolved [2].

4) NEURONS BEHAVIOR ANALYSIS IN HIDDEN LAYERS PROBLEM

The difficulty of analyzing neurons behavior in hidden layers is another identified obstacle of ANN implementation to PR.

That is, it's difficult in analyzing the performance of some neurons in hidden layers or it is difficult to interpret the behavior of some neurons in the hidden layers. However, if the performance of neurons' behavior in hidden layers can be successfully analyzed, it will help in identifying false classification. The success analysis of neurons behavior in hidden layers can reveal information about noise in the data or corrupt data [71]–[74]. Therefore, more investigation is requiring addressing the problem of neuron behavior analysis in the hidden layers for better success in ANN techniques for PR.

5) HANDLING OF LARGE DATA PROBLEM

A connected feed-forward neural network (FFNN) that succeeded in character recognition always have problems with *large variables* to handle. That is, FFNN has the difficulty of handling large data variables images [75], or FFNN has a challenge of handling many variables with spectral representations of spoken words.

For example, a network that has its 1st layer connected to 10 hidden units can contain several weights (up to 1,000) that may lead to overfitting issues if training data is large or limited [40]. However, overfitting issues cannot be necessary attributed as a general problem in ANN. Similarly, the network memory requirement for many weights can prevent certain hardware implementations and causes variations in the inputted objects.

Only some category of FFNNs that can be applied to PR [76]–[78], that is, not all classes of FFNN models can be used for the PR task. The term feedforward indicates no response to input patterns or data. Just as human being learns from tasks, similarly, neural models can learn or acquire knowledge from training by way of feedback to inputted data [79]–[81]. Usually, the feedback helps in recreating input patterns.

Also, feedback makes input patterns to be free from mistakes or error or produces an error at a negligible level; by so doing it expands and improves performance of the NNs [79]–[81]. Such an improved network is called the auto-associative neural network, which can be challenging to construct [82], [83]. Hence, more study is requiring addressing the challenging area of error-free network construction.

6) LOCAL MINIMA PROBLEM

A traditional algorithm like backpropagation is used for training of NN, but this has a problem of local minima [5]. In performing a task, auto-associative, neural networks use back-propagation algorithms [84], [85]. Backpropagation is a technique utilized in ANNs to compute a gradient that is required in the calculation of weights found or used in a network. However, the key problem related to back-propagation paradigms is its tendency to local minima [86], that is, the value becomes small instead of getting large. Hence, a back-propagation network or paradigms require further research for a more significant result.

Nature inspires algorithm can be used for training of NN. This nature inspiration algorithm is helpful in determining global optima. An algorithm is utilized for the learning process. For instance, the genetic algorithm executes a parallel search that can enhance computational speed.

Likewise, tabu search enables a flexible and robust memory [5]. Optimization of the ant colony (OAC) can be applied to weight optimization. An enhanced cuckoo search enable flexibility to a parameter to improve the speed and accuracy.

7) LONG TRAINING TIME ISSUES

Most training of NN towards PR can consume time, and their structure is complex to design, especially designing a deep NN. Therefore, an investigation is requiring to addressed time consuming and structural design problem aspect of ANN that focuses on PR.

8) HIGH COMPUTATIONAL COST CHALLENGE

Cost is highly significant to any profit and non- profit-making organization. Currently, the cost of cognitive computational intelligence like ANN models, robot or ML paradigm is enormous. Hence, research is requiring addressing the computational cost challenges of ANN models [5].

9) WEIGHT ADJUSTMENT PROBLEM

Currently, the necessity of weight adjustment in ANN is a combinatorial problem, and to discover the desired output; there is a need for weight optimization. Methods of weight adjustment, such as backpropagation (BP) and some non-iterative methods required in weight adjustment.

Usually, the method of NNs training can be based on the initial set of parameters, weight, bias, and algorithm learning rate monitoring. It begins its leaning using some initial value and updates of weight in each iteration. The weight adjustment issue and feature made NN less fitting to data mining classification.

These unresolved problems over the years have made ANN application to encounter retardation in PR tasks. However, ANN technology is still evolving as more research can lead to a significant solution in those areas. Although many problems and hindrances exist in ANNs application to PR, nevertheless, the potentialities of ANNs are numerous and powerful, so its advantages over other alternative models are quite huge and divergence [87]–[89]. Future work is requiring to analysis some challenges in weight adjustment, such as gradient vanishing problem.

10) DECODING THE PATTERNS OF THE HUMAN BRAIN PROBLEM

Another global un-addressed question in machine learning (ML) and neuroscience is whether computers can decode the patterns of the human brain [90]. This novel idea is a potential area for future study of ANN application to PR. Importantly decoding shows that information is latent in the patterns of NN activity, that means, information is existing in NN but not yet developed. Notwithstanding, researchers always draw a further inference, that is when decodable information exists, then there is a strong proof that the information is represented by the patterns of activity utilized as the basis for the decoding. Analyses of brain activity patterns can help in revealing what an individual is thinking, seeing, remembering or focusing on.

Thus, ANN models can be used to investigate how the brain encodes complex and multicomplex abstract semantic information or visual scenes. Such feats of ''brain reading, and mind-reading if successful it will be an impressive research breakthrough.

B. SPECIFIC ANN APPLICATION PROBLEMS TO PR

Now, let discuss some specific ANN application or PR problems which includes;

1) ARBITRARY ORIENTATION COMPLEXITY

Currently, there are complexities in the PR of an object with whimsical or arbitrary orientation even with ANN. That is, some objects have an unknown path that cannot be accurately resolved due to their directional position. The unknown directional position makes it difficult for ANN to recognize their pattern. In this context, ''arbitrary'' means an unknown path while ''orientation'' means a directional position.

Most existing research using ANN has devoted to recognizing horizontal and near-horizontal texts but ignoring the area of the object with arbitrary orientation. With the rapid growth of smartphones and practical vision systems, the issue of text recognition in natural scenes is becoming critical, yet a challenging task in human-computer interaction [61].

The natural scene involves the process in which an agent like a human being visualizes and interprets environment things in the place where event or action occurs (e.g., living rooms busy streets, meadows, etc.). Therefore, research is requiring in the area of computational intelligent to address complexities in the PR of an object with arbitrary orientation in any environment.

2) OBJECT CLASSIFICATION CHALLENGES

ANN is one of the widely applied techniques for classification. Classification referred to as the process of the grouping of things, items, or objects by shared characteristics, features, and qualities. Thus, ANN can carry out the task of PR of an object. However, the challenges of PR come when many features are alike in objects. These many features can cause difficulty in differentiating one object from another object.

Hence, research is requiring in this direction of differentiating an object of many features that are alike from one another.

In practice, the backpropagation NN (BPNN) can be utilized as a success tool for dataset classification of such an object with likeness features. The dataset classification is achieved in an object of likeness features with a fitting combination of training, learning, and transfer functions [62]. Though BPNN can be applied to achieved success, nevertheless, the applied BPNN method performance is not significant enough to accurately address the problem of differentiating an object with likeness features from one another.

3) OBJECT LOCATION DIFFICULTY

An object analysis or location is one of the critical roles of learning in ANN. Object location is a problem that involves the extraction and processing of crucial information from the complex and uncertain object data in limited time [63]. In this way, the ANN application can help in addressing the human difficulty in an object location in a short period. Therefore, research is requiring addressing this difficulty of object location in the hidden environment for a significant result that will improve an object location task.

4) RULE AND TEMPLATE MATCHING

Furthermore, ANN recognition of pattern challenge occurs in rule and template matching kinds of paradigm [91]–[93]. Both template and rule matching algorithms require learning capability, especially in the stock market prediction via PR of market price determinant variables [94]–[97]. Hence, further research is urgently requiring in the area of template and rule matching algorithms for a better outcome.

5) SPEECH RECOGNITION ACCURACY CHALLENGES

Speech recognition (SR) accuracy is still a challenge with ANN. Also, over the years in research development, it is found that ''accuracy of ANN in SR continues to be one key research challenges, for example, in the areas of speaking and language variability, speech domain, noise recognition, and vocabulary volume or size needed or used in speech [98]–[101]. Although SR is a difficult problem to handle, however, an improvement in any ''individual SR accuracy'' can be an improvement in the entire system performance. Hence, further study is requiring addressing the challenging area of speech recognition.

6) IMAGE RESTORATION PROBLEMS

In performing PR tasks, ANN has image restoration problems that bring noise, especially from a compressed image [102]. An image restoration without noise is required in pattern recognition using ANN. Therefore, more research is necessary for the area of image restoration problem with ANN application by interested researchers.

7) IDENTICAL IMAGE RECOGNITION CHALLENGE

When one considers two images that is comparably identical, one of the images may not be accurately detected. An example, an ANN model like a deep neural network (DNN) could recognize or detect the first identical image rightly but may not recognize the second image.

Current discovery by some Google researchers at Universities in New York and Montreal finds that image restoration defect or problem is almost in every DNN [103]. Hence, this aspect of an image restoration defect needs further research to enable a meaningful solution and widespread application.

8) SURFACE MATERIAL RECOGNITION ISSUE

Material recognition has been a long-standing issue with the ANN. Thus, materials in image recognition have become a challenging task for ANN [104]. Real-world materials have surface texture, lighting, geometry, and clutter, that combined to make PR remains difficult for ANN during application. Therefore, surface material recognition using ANN is another prospective area of research for an interested researcher.

9) RECOGNIZING PLANNED ATTACKS

In ANN application to PR on the natural incidence or otherwise, there is a need to extend its functionality to global phenomena. An example, an ANN approach is requiring to recognizing router sign flooding attacks in the next generation of IPv6 networks [105]. In general, NNs have issues related to feature representation, modularity, architecture selection, scaling, and learning speed [106]. Hence, more study is requiring in addressing these areas of ANN application to PR in natural incidence for a tangible solution.

V. NEW AND CURRENT ANN MODELS APPLICATION TO PR

Currently, feedforward NNs like Single-layer perceptron (SLP), Multilayer perceptron (MLP), Radial basis function network (RBFN), generative adversarial networks (GANs), sparse auto-encoder (SAE), deep belief network (DBN), reservoir computing, transformer models are performing great in PR. Likewise, the use of some class of feed backward neural network (FBNN) like RBM, SOM, and Hopfield networks, competitive networks, deep neural network (DNN), and hybrid models are effective in PR.

Also, current hot spots in PR using ANN is; the use of convolutional neural networks (CNN), recurrent neural networks (RNNs). Other new areas include probabilistic neural network (PNN), time-delay neural network (TDNN), and Kohonen self-organizing NN or self-organizing map (SOM). These ANN models for PR can be described as follow;

A. GAN DEEP LEARNING FOR PATTERN RECOGNITION

Nowadays, an intelligent diagnosis paradigm using generative adversarial learning deep neural networks (GALDNNs) and their application to PR in planetary gearbox fault is becoming popular. Previously, many results become suboptimal from researches, texts obtained with Markov models were dull and predictable.

Likewise, variational autoencoder's images were blurry and lacked varieties [107]–[109]. However, currently, generative adversarial networks (GANs), is applied to trained with few data which could provide pattern recognition in missing data [110], [111]. GAN is applying to the future prediction of the frame and in high dimensional spaces such as 3D modeling [112]. Generative adversarial networks (GANs) conceptual approach is an example of unsupervised learning because it does not require to be feed with expert knowledge. The concept of generating samples is a given dataset that does not require human intervention. GAN flow of data can be represented as in Fig. 7.

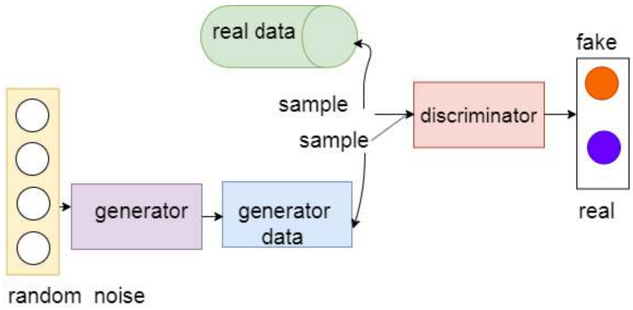


FIGURE 7. The architecture of a typical GAN.

The reason behind GAN success in PR is that the structure consists of two NNs that relatively content in a zero-sum game framework, that is, the generator model and the discriminator model.

The two models are trained together in a zero-sum game, adversarial or enmity, until the discriminator model is deceived about half the time, meaning that the generator model is generating plausible (fake) looking samples. Generator model trained to generate a new image, while the discriminator model tries to categorize samples as either real (i.e., from domain) or fake (i.e., generated). GAN advantage of having both a discriminator and generator makes it powerful in its application to PR. Thus, GAN is a deep NN architecture composed of two nets conflicting one another (the word adversarial) [113]–[116].

GANs are performing well and rapidly changing a field during the application, it delivers the promise of generative models in their ability to produce realistic samples across a range of problem domains, specifically in image-to-image translation and PR tasks such as translating photos of the night today, or from one season to another season like autumn to spring and in generating photorealistic photos of events, objects, items, and people that even human beings cannot know are fake.

B. SAE FOR PATTERN RECOGNITION

An autoencoder NN is an example of an unsupervised ML algorithm that uses backpropagation in setting target values or outputs to be equal to the inputs. Autoencoders application includes the conversion of any white and black picture into a color picture. Likewise, autoencoders are used to reduce the size of our inputs into a smaller representation. If anyone needs the original data, they can reconstruct it from the compressed data.

Recently, sparse auto-encoder (SAE) is applying in PR to increase generalization performance in human motion [117]. An autoencoder is a kind of ANN utilized to learn data coding efficiency in an unsupervised way. The purpose of an autoencoder is to understand an encoding (representation) in a dataset, particularly for the reduction of dimension through

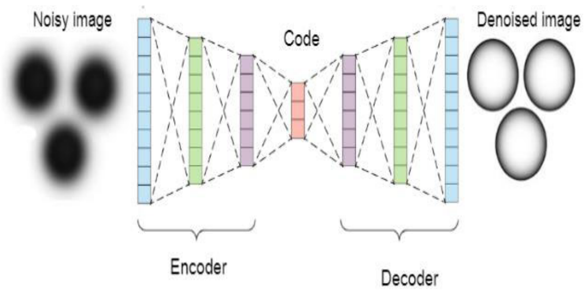


FIGURE 8. The architecture of a typical SAE.

training of network in ignoring the noise (signal). An architecture of a typical SAE is shown in Fig. 8.

Fig. 8 represent the trained SAE that reconstructs the original input image from the noisy image version. SAE extracts only the features of an image and produces the output by eliminating any disturbance or unnecessary noise in the system. In Fig. 8 the encoder layer converts the noisy input image as a compressed presentation in a lessen dimension.

The code layer, also known as the "bottleneck," of the network presents the compressed image inputted into the decoder. The decoder layer translates the encoder image (noisy image) back to its original dimension (denoised image). Indeed, autoencoders learned automatically from examples.

Therefore, it is easy to train an algorithm that can perform significantly on a specific input. In 2014, the study by Liu, & Taniguchi [118], applied deep sparse autoencoder that extracted low dimensional features, which represent the characteristics of individual motion efficiently, from data of human action or motion with high dimensional. Recently, the SAE ideal has become popular for generative models learning of data [119]. Most robust ANN models in the year 2010 include SAE in DNNs [120], [121].

C. DBN FOR PATTERN RECOGNITION

Recent studies have shown that deep belief nets (DBNs) are applying to various problems ranging from speech recognition, language understanding to image, and audio classification problems [122]. A typical DBN, with circles and arrows showing the flow of information among restricted Boltzmann machine (RBM), can be represented as in Fig. 9.

DBN is formed from stacked restricted Boltzmann machines (RBM), as shown in Fig. 9. Each RBM is a two-level model, having visible units of layer and hidden units of a layer. In machine learning (ML), DBN is a generative graphical model, or a class of DNN, comprised of multiple layers of hidden units (latent variables), with interconnections between the layers. However, the connections are not between the units in each layer. DBNs are a subset of DNNs; the difference between the two is mostly on how they are trained to perform the task at hand. The pretraining of DBN proves to be beneficial if the training set is small. Currently, DBN is applying for natural language understanding with significant success [123].

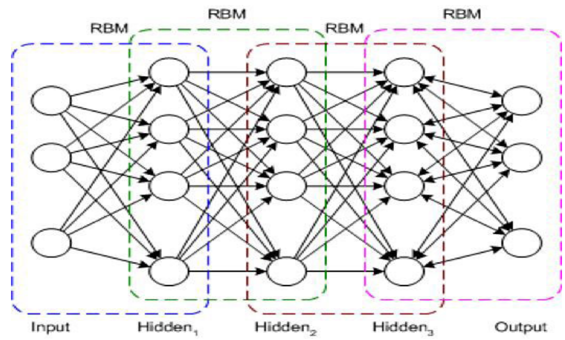


FIGURE 9. A typical DBN, showing the flow of information among RBM.

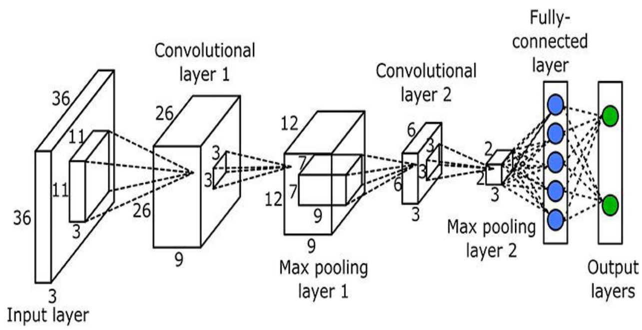


FIGURE 10. An architecture of CNN.

D. CNN MODELS FOR PATTERN RECOGNITION

Recently, convolutional neural networks (CNN) have wide applications in video and image recognition, processing of natural language [124], and recommender systems [125]. Also, CNN has a popular application in speech recognition and distant speech recognition that has produced a better result than the DNN. CNN performed excellently in image data, PR, classification, and regression. However, generally, CNN's function well with spatial data [126] and its application have demonstrated the most accurate results in solving real-life problems [127]. The architecture of CNN can be represented in Fig. 10.

As shown in Fig. 10, CNN input is conventionally two-dimensional, a matrix or field. Also, it can be modified into one-dimensional, thereby making it build an internal representation of one-dimensional order. One remarkable advantage of CNN is that it can give better accuracy and enhance system performance due to its distinctive features like local connectivity and shared weights.

CNN is remarkably better in achieving success than other deep learning approaches in applications, especially in their application to natural language processing and computer vision because it can reduce most traditional problems [127].

CNN's generally do not perform excellently when input data depend on each other in a sequential pattern. CNN's can achieve state-of-the-art output on challenges like text or document classification used in sentiment analysis and related issues. Currently, CNN is useful in text data, time-series data, and sequence input data.

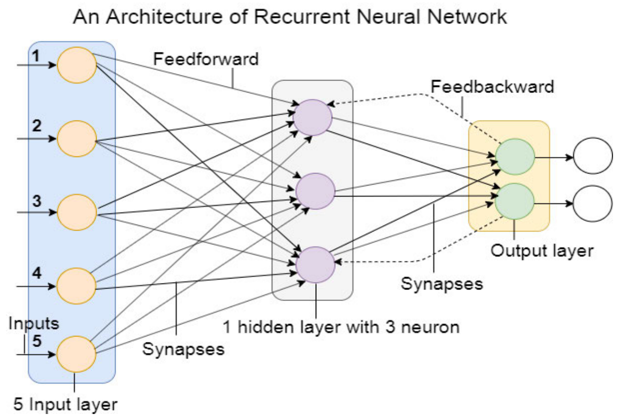


FIGURE 11. A typical structure of RNN.

A significant advantage of the CNN over conventional approaches is its ability to concurrently extract features, then reduce data dimensionality, and capability to classify in one network structure. Additionally, the CNN technique requires only minimal image preprocessing due to the robust ability of CNN to noise during images acquisition.

E. RNN MODELS FOR PATTERN RECOGNITION

Recurrent neural networks (RNNs) are useful tools for PR and can be suited for powerful sequence learning tasks [128]. RNNs have proven to be great pattern recognizer and predictive engines, especially in a task involving sequence Machine learning like speech recognition. RNNs are superb in PR and predictive accuracy, and sometimes no algorithm can compete relatively with RNN. RNNs has feedback loops in their recurrent layer. This helps them keep information in 'memory' for an extended period.

However, it can be challenging training a standard RNNs to resolve problems that need learning for a long period of temporal dependency. This is because of the loss function gradient decays exponentially over time (which is referred to as gradient vanishing problem).

Long short-term memory (LSTM) networks are a kind of RNN that utilizes specific units together with standard units. A 'memory cell' is a component of LSTM units that can keep information in memory for a long time. LSTMs are often called to as fancy RNNs. While Vanilla RNNs have no cell state. Meanwhile, they have hidden states only, and the hidden states function as RNNs memory. A structure of RNN can be depicted in Fig. 11.

Arrows in Fig. 11 demonstrating the inter-relationship amongst FFNN components. RNN has achieved remarkable success in sequence learning problems. The success achieved in terms of application to PR includes accuracy in speech recognition and timing. One of the studies of applying RNN is the recent work by Wu, Ding, and Huang [129], which proposes a novel method for fault prognosis on the deterioration sequence of equipment. The proposed method tested for performance using the health data of aircraft turbofan engines. Result demonstrates that the RNN can perform well in one-step, long-term, or in remaining useful life prediction tasks.

Currently, CNN is performing remarkably well in PR; its design applicable to facial expression can be superb. During application, the input plane detects, recognizes, validates the image of the character that required normalization. Then the layer unit accepts inputs from the units in the preceding layer [130]. CNN's uses the variability of MLP constructed to handle minimal preliminary processing [131].

CNN's is also named shift variant (SVANNs), or space invariant ANNs (SIANNs) based on related weights design and invariance behavior. A major advantage of CNN's is its ease of use during implementation.

Mamuda and Sathasivam [132] proposed ANN in medical examination. The outcome reveals the MSE of the high correlation between target and output predicted. While, Lee KY, Chung, and Hwang [133] investigated medical challenge with NN, the result shows NN could predict high variability in medical patterns. Likewise, research by Dicky et al., [134] demonstrates NN recognition of distinct in Eye iris.

Cheng and Sutariya [135] paper on drug discovery using NN demonstrate learning in a PR system that can diagnose disease and make a prediction of disease. The study implied NN that handles challenges in complex PR. Cheng and Sutariya concluded that ANNs' application to drug discovery has a prospect for future research due to its usefulness in medical classification.

The study by Agatonovic and Beresford [136] applied NN to a pharmaceutical challenge as an alternative to the conventional surface approach, and interestingly, the result proves useful with the supervised NN.

Christodoulou and Pattichis, [137] applied NN to a medical the issue, and the result demonstrates 97.6% accuracy compared to the 95.3% success from the alternative a statistical model with the same dataset. The unsupervised PR enable knowledge on the category of motor unit action potentials (MUAPs) in electromyographic (EMG) for recognition of neuromuscular disorders.

In recent times "instant physician" trained an auto-associative memory NN that stores various medical items like patient symptom's data, patient's test records, and treatment cases [138], [139]. After the training, the NN produced input containing a set of symptoms. Systematically, the result shows that NN finds a full pattern that constitutes the "best" treatment and test for any particular medical issue.

In 2017, a study by Alexey *et al.* [140] trained CNN's that generate images from objects taken concerning the object's color, viewpoint, and style. The result reveals that supervised trained CNN could be used for PR tasks and for effectuating images with given features like lighting information, viewpoint, and high-level style. A network trained for a generative task can learn to generate training samples and learn generic presentation. Generic implicit representation aids a network to smoothly morph across various object views or object instances in a meaningful image.

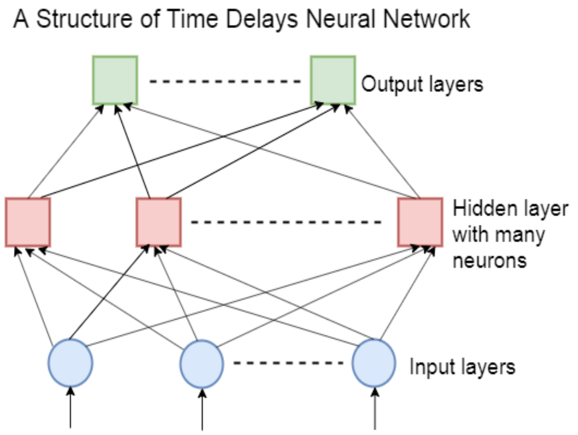


FIGURE 12. A structure of time delays neural network (TDNN).

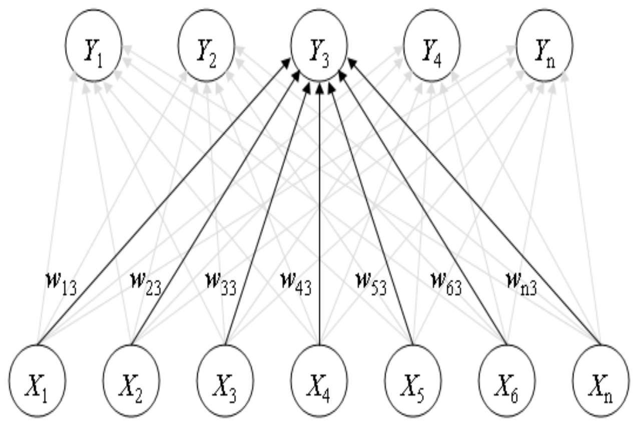


FIGURE 13. An architecture of a SOM network.

F. PNN MODELS FOR PATTERN RECOGNITION

A probabilistic neural network (PNN) is a kind of four-layer feedforward NN. The four-layers are the input, the hidden, pattern/summation, and the output. In the parent probability distribution function (PDF) of the PNN algorithm, each class is approximated by a non-parametric function and a Parzen Window. When the PDF of each class is in use, the new input class probability can be estimated, and Bayes' rule is applied to allocate the highest class of the posterior probability [119]. PNN was a derivate of both the Bayesian network (BN) [119] and the Kernel Fisher discriminant analysis (a statistical algorithm) [141]. It is commonly useful in PR and classification.

G. TDNN MODELS FOR PATTERN RECOGNITION

A time-delay neural network (TDNN) is a typical feedforward architecture that recognizes features in sequential data not depending on arrangement position. In attaining time-shift invariance, delays are included in the input to analyze multiple data points. This analysis comprises typically a more extensive PR system [142], [143]. TDNN is another kind of approach for data classification that gained recognition for many years. It can perform remarkably well on time series and widely applied to speech recognition, image sequence analysis, and stock market prediction. A typical structure of TDNN can be depicted as in Fig. 12.

Fig. 12 represents a time delay NN whose input consists of few tap delays lines.

Furthermore, it can be more efficient in the implementation of embedded systems when compared to the decision trees method. Examples of embedded systems include MP3 players (referred to as MPEG audio layer 3), i.e., MPEG (moving picture experts' group), digital cameras, video game consoles, mobile phones, GPS (global positioning system) and DVD (digital video disc) players. Household embedded systems are washing machines, dishwashers, and microwave ovens. Thus, embedded systems are mostly found in industrial, medical, consumer, automotive, police, military, and commercial applications to provide features, flexibility, and efficiency.

H. SOM FOR PATTERN RECOGNITION

The Kohonen self-organizing NN or self-organizing map (SOM) is a useful tool for PR [144]–[146], it is an effective tool for multidimensional data analysis [147]. An example self-organizing network having five cluster units, the five cluster units are organized in a linear array together with Yi, which have seven input units, X_i, which is shown in Fig. 13.

Each weight "w" of the SOM network is a representative of a particular input. Input patterns are revealed to all neurons concurrently. SOM is a kind of ANN that uses unsupervised learning to train. It creates discretized low-dimensional representation of the space input of the samples training that is referred to as a map, which is a method of dimensional reduction. Interestingly, SOM networks may be either unsupervised or supervised.

Unsupervised learning involves modifying the weights of a NN with no consideration for specifying the output for specific input patterns. The advantage is to enable the network to find its solution, thus making it more effective and efficient with pattern connection. The main disadvantage is that users or other programs need to think of interpreting the output.

I. TRANSFORMER MODEL FOR PATTERN RECOGNITION

An ANN model has proven to be effective for common natural language processing tasks; the model is called a Transformer which makes use of different mechanisms and methods. Transformers are a type of NN architecture that has been gaining popularity. A transformer is designed to resolve sequence-to-sequence recognition tasks while handling long-distance dependencies between input and output with attention and recurrence. The task that transforms an input sequence to an output sequence includes text-to-speech transformation, speech recognition, etc. Recently, transformers were used by OpenAL in their language models and used by DeepMind for AlphaStar [148], [149], [151], [152].

J. RBM MODEL FOR PATTERN RECOGNITION

A restricted Boltzmann machine (RBM) is a generative stochastic ANN that could learn a probability distribution in its set of inputs. They have found applicable in dimension reduction, collaborative filtering, modeling, classification, feature learning, images, quantum mechanics, etc. They can be trained in unsupervised or supervised ways; it depends on the task. By contrast, unrestricted Boltzmann machines can have connections in hidden units. The restriction enables efficient training paradigms than the general class of Boltzmann machines, particularly the gradient-based contrastive divergence paradigm. RBM has attracted attention as building blocks for the multi-layer learning systems referred to as deep belief networks (DBN), variants and extensions of RBMs have found application in a wide range of PR tasks.

K. HOPFIELD NETWORK APPLICATION

Most NN models fit into one business area of resource scheduling, allocation, and management. ANN model applied in business include financial and accounting analysis. NNs like Hopfield-Tank network could be applied to scheduling and optimization. An experiment has shown that NN effectively learned pattern features and recognized patterns correctly [2]. Nowadays, in the stock market, ANN is rapidly applying to technical trade analysis in the stock business and related area and security exchange [2].

L. DEEP DIRECT REINFORCEMENT LEARNING

Recently, the research by Deng *et al.* [153] proposed a deep direct reinforcement learning (DDRL) to financial sign depiction and trading, and the result proves effective to stock market business. A similar investigation was conducted by Göçken *et al.* [154] with the application of NN to price pattern of the stock market. The outcome demonstrates PR with 3.38 prediction for Mean Average Percentage Error (MAPE). In the year 2016, an investigation by Qiu, Song, and, Akagi [155] on the stock market using ANN, shows the success of PR via prediction of returns in the stock market.

M. HYBRID NETWORK MODEL APPLICATION

A study by Tan [156] introduces a hybrid financial system that applied statistical models, genetic algorithms (GAs), chaos theory, and, ANNs. The result shows that ANN was more efficient in intruder detection regarding PR compared to other methods using the same dataset.

System load forecasting plays a significant role in the energy management system (EMS). Previously, ANN applied to electric power system load forecasting. In recent times load forecasting was one area of study in electrical engineering that requires research attention for improvement. To improve the precision of EMS short-term load forecasting, Dai, and Wang, [157] proposed ANN to forecasting short-term loading system.

Thus, the applied ANN significantly influences the operation, controlling, and planning of the EMS.

Likewise, trained ANNs have patterns with forecasting accuracy. More also, load forecasting is useful in safety, planning, economic, and operation of EMS. Furthermore, ANN is found useful in start-up and shutting down schedules of generating units, load-management, and maintenance of power systems and planning.

N. RESERVOIR COMPUTING

Reservoir computing is a framework for cognitive computation that may be taken as an extension of ANNs [158], [159]. Normally an input is fed into a fixed (random) dynamical system termed a reservoir and the dynamics of the reservoir map the input to a higher dimension. Then a simple readout mechanism is trained to read the state of the reservoir and map it to the desired output. The main benefit is that training is performed only at the readout stage and the reservoir is fixed. Examples of reservoir computing are liquid-state machines (LSMs) and echo state networks (ESNs).

The word liquid from liquid-state machines comes from the analogy of dropping a stone into a stable body of a liquid like water. The falling stone can cause ripples in the liquid. The input, that is, a motion of the falling stone has been converted into a Spatio-temporal pattern of liquid displacement (ripples).

LSM is a kind of NN that consists of a large collection of units (neurons or nodes). Each of the nodes receives time-varying input from the external sources i.e. the inputs as well as from other nodes. Nodes are randomly connected. The recurrent nature of the connections turns the time-varying input into a Spatio-temporal pattern of activations in the network nodes. The Spatio-temporal patterns of activation are read out by linear discriminant units. LSMs can be a way to explain the function of brains.

LSMs are said to be an improvement over the theory of ANNs because its circuits are not hardcoded to perform a specific task. Also, continuous-time inputs are resolved ''naturally''. More also, computations process on various time scales can be done using the same network. Furthermore, the same network can perform multiple computations.

In the case of echo state networks (ESNs). ESN is a recurrent NN with a sparsely connected hidden layer about 1% connectivity. The connectivity and weights of hidden neurons are fixed and randomly assigned. The weights of output units are learned so that the network can reproduce specific temporal patterns. The interest of ESN is that though its behavior is non-linear, the weights that are modified at training are for the nodes that connect the hidden units to output units.

Hence, the error function is quadratic with the parameter vector, which can be differentiated easily into a linear system. Recently, a research team from the University of Michigan implemented the reservoir computing rules in a chip and showed its performance in a speech recognition task.

O. MLNN APPLICATION

ANN is a subset of machine learning (ML), machine learning models, methods, or learnings can be of two types that are unsupervised and supervised learnings. Machine learning neural network (MLNN) is a class of supervised learning (SL), where the model first acquires training from the training dataset and after learning it then classifies image tested using the acquired knowledge. Neural networks (NNs) are a specific set of algorithms that have revolutionized machine learning, and the current so-called deep neural networks have proven to work quite well in PR.

Machine learning has proven to be useful in many applications to recognizing hidden patterns such as in weather and satellite communication. The research by Abhishek et al., [160] utilizes ANN to rainfall and weather based on analyses of rain attenuation factors to solve rainfall and weather attenuation problem of signals unavailable for satellite communication.

Abhishek team designed a rain attenuation model with different NN topologies. Their result demonstrates that the applied ANN (Machine learning) shows differences in the pattern in hidden neurons of the NN topologies. The approach demonstrates an effective way of predicting rain attenuation using ANN in microwave satellite. Finally, Abhishek et al., the study pointed out that MLNN or MLANN algorithms perform than SLANN algorithms in terms of PR. Then, the higher the input data lesser the MSE of trained ANN.

In related work on protecting of wireless sensor networks against sensitivity to failure links and nodes, a recent study by Swain, Khilar and Dash, [161] discovered a multiple fault recognition framework in sensor system (SS) via hybrid metaheuristic NN. The result showed that hybrid metaheuristic NN, FFNN model can use PR technique to diagnose composite faults like transient faults for sensor nodes and links, intermittent, soft and hard permanent faults.

Currently, a hybrid feedforward NN have shown significant success in PR as applied to sediment load estimation [162]. In a separate study on evaporation, Ghorbani *et al.* [163] and Qasem *et al.* [164] applied ANN to identify the trend of pattern in monthly evaporation in the arid and humid climates. Both studies on evaporation demonstrate that ANN provided unique trends for evaporation modeling.

Likewise, research by Ahmadi *et al.* [165] on how dynamic viscosity affects the transferring of heat and flow of fluids using ANN to identify patterns revealed remarkable success.

In recent time, ANN applied to PR in rainfall modeling [166], and sugarcane growth [167] using meteorological parameters, which shows effective performance. Research by Abidoye and Chan [168] used AI to a valuation of the property. The research enumerated the urgency for shifting from conventional practices management to applied ANN in property valuation, especially through pattern identification. Crowdsourcing and active learning (AL) are an effective, efficient, and successful way to accomplish PR.

In 2014, researchers like Vijayanarasimhan, and Grauman, [169] presented live learning of object detectors, and the system autonomously refines its models by actively requesting crowdsourcing. The result has shown that heighten recursion profundity could aid the accomplishment of an image in the absence of new parameters for more convolutions.

In an attempt to achieve an abusive success in forecasting the load power system for the short-term. The study by Dai and Wang [157] applied MLFFNN with the BP learning algorithm trained samples. The results from the test show effective forecasting for load in a short-term power system.

P. DEEP RECURSIVE NETWORK APPLICATION

In 2016, Kim *et al.* [55] proposed image super-resolution (ISR) by applying deeply recursive CN (DRCN) to simplify training for PR. The investigation used extensions like skip connection and supervision recursive to show that DRCN has DR layer as much as 16 recursions, reuse weight parameters that exploit a large image context. The DRCN method can be applying to other image restoration challenges like artifact removal and denoising.

Q. DEEP NEURAL NETWORK EMERGENCE

Nowadays, deep learning researchers are applying several methods for enhancing visual task. Various domain in deep learning (DL) have been studied to improve PR, for instance, the use of RNN and DL in video production and an end-to-end task, such as upgrading, structure, and parameters specification [156].

Recently, deep NNs (DNNs) shows impressive performance in PR tasks. In the year 2017, work by Dong *et al.* [170] found neurons in DNNs did not recognize semantic objects or components but recognize objects or components as periodic base reinforcement. Similarly, no active organization in DNN visual study codes abstraction, because the designs of adversarial objects are large and therefore not uniform with real images. Though, homogeneity exists in DNN visual description, and visual appearance, but, distinct from past discoveries. More result from Dong et al. work shows that neurons understand powerful fundamentals and neurons representations ally across adversarial and real images.

Over time DNNs have achieved success in many PR tasks like in image classification. In 2017, the research by Liu *et al.* [171] on deep recursions introduced a visual analytics technique to understand further, diagnose, and refine DNN. However, the advancement of quality DNN designs usually depends on trial by error, until explicit knowledge of how DNN operate. FFNN technique to generic scheme to train a network in achieving navigation tasks are still new. Kelchtermans, and Tuytelaars [172], deployed FFNN an unmanned aerial vehicle (UAV) to proffer a general principle that applies control for navigation tasks.

Some data can be modeled normally without observing the order of structure [173]–[175]. A novel investigation by Ryan *et al.* [176] provides a robust for undetermined data order to image recognition. Also, the design uses a nested stick-breaking algorithm that enables unbounded depth and width for data to remain anywhere at the node and substituted infinitely.

More also, the novel investigation uses many replacement and presentation that makes data to be by internal nodes without a problem. The proposed technique can be applying to orderly clustering images and text data modeling.

Nowadays, in the manufacturing industry, NNs are applying to recognize electronic malfunctions. A study by Murphy and Kagle, [177] applied NN software to process control for recognizing an electronic circuit board in the manufacturing industry. The result showed that the NN considerably mitigates the time needed to construct a PR system.

Likewise, the research by Vellanki, and Dagli, [178] applied NN with knowledge-based systems to PR in the operation of the circuit board assembly task. The result demonstrates that realtime testing of component recognition is effective with the NN model.

An investigation by Jack *et al.*, [179] uses FFNNs to resolve inverse kinematics difficulty in robotics in three distinct cases. The result proves that NNs can be an alternate technique for inverse kinematics estimation, because of NNs advantage of fault-tolerant and high-speed to inverse kinematics problem in robotics.

Tool wear management is a key issue related to many material removals or elimination processes. Recently, the paper by D'Addona, Ullah, and Matarazzo [180] reveals that both ANN and deoxyribonucleic acid (DNA) based computing can be useful in recognizing tool wear. As a result, it predicts tool wear from a set of tool-wear images. Also, more result from D'Addona et al. studied reveals that DNA-based computing (DBC) can recognize patterns or likeness and dissimilarity in processed images. Excitedly, the study can further be used to address other complex problems using a combination of DBC and ANN. That is, in a situation whereby both PR and prediction are two severe computing problems that required to be handled concurrently.

Designing new products is vital to business success and sustainability in the manufacturing industry. In the past years market, competitiveness becomes a key stimulating factor for designing good products. For example, in the sanitary ware industrial sector, the design of new sanitary ware products requires a long amount of time. Likewise, it requires a cost-effective process because of planning, marketing strategies, designing, and testing of a new product. That is, it requires strategies that can provide a solution to the problem of reducing product designing repetitive time, improving production and quality.

To solve these problems, the study by Brahim, Smith, and Bidanda, [181] applied a NN to improved product design in many manufacturing industries, particularly in sanitary ware manufacturing. The result shows that the NN model on product design and testing was effective.

Nowadays, there is research in engineering, that focuses on predicting the time a component or system will malfunction [182]. The progressive knowledge in AI has helped the growth of ANN to PR, and recently, one can analyze the diverse uses of ANN. ANN has several uses as it provides solutions to simple, complex, and multi-complex challenges [183]–[186]. ANNs are excellent recognizers of patterns and robust classifiers [187]–[190] that have the capability to generalized when deciding on inaccurate data input [191]–[194].

More also, recent researches in the AI areas of deep learning (DL), reinforcement learning (RL), and combine (DL) and (RL) seems promising in transforming the future of ANN [195] to greatness. These wonderful achievements in ANN can be observed from the unique characteristics of ANNs. These distinct characteristics include massively parallel structure, ability to store experimental knowledge, high interconnectivity to networks, self-organizer, and the collateral plethora of human's visual system characteristics. There are lots of methods involved in ANN PR. However, researchers in the field need to identify problems in ANN application to PR and focus on those areas for greater success.

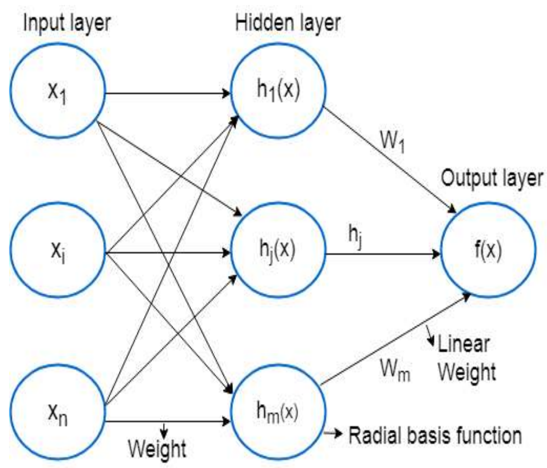


FIGURE 14. **The architecture of a radial basis function network (RBFN).**

R. RADIAL BASIS FUNCTION NETWORKS

An RBFN is a type of ANN which utilizes RBF as activation functions (AF) [196]. RBFN accepts input vector as input each with divergent parameter. Network output is neuron parameters and a linear combination of RBFs inputs. RBFNs input modeled as real numbers vector. RBFNs in its simplest form in three layers; input layer (IL), a hidden layer (HL) with a nonlinear AF, and linear output (LO).

Another performing a PR network is RNNs, which is a network that has at least one feedback relation or connection. Thus, RNNs is dissimilar to RBFN, RNNs can have hidden units or may not have hidden units depending on design usage.

The structure of an RBFN can be represented as in Fig. 14. where,

$$f(x) = \sum_{j=1}^{m} w_j h_j(x) \tag{1}$$

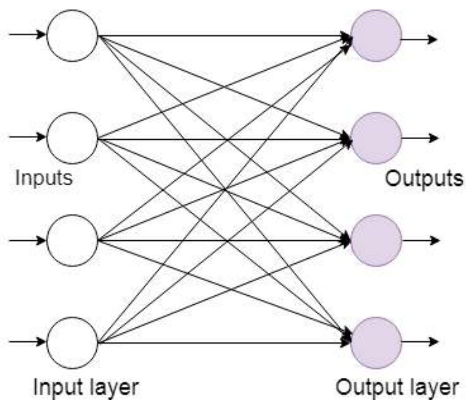


FIGURE 15. A single-layer NN.

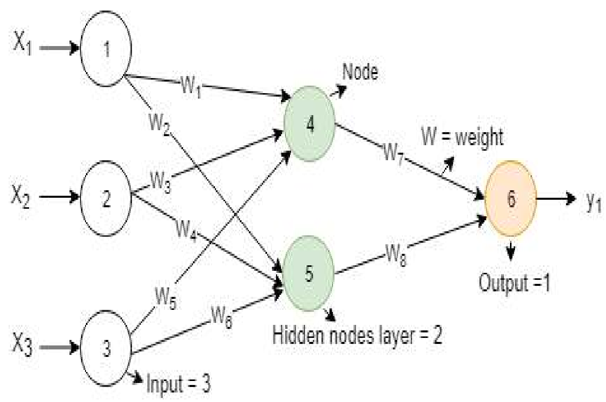


FIGURE 16. The architecture of multilayer feedforward neural network.

As shown in Fig. 14, RBFNs has diverse uses such as function approximation (FA) [197], prediction [198], classification [199], control [200] etc.

S. SINGLE-LAYER NEURAL NETWORK

Single-layer neural networks (SLNNs) are a class of feedforward neural networks (FFNNs) that are useful in PR. Single-layer FFNNs, (SLFFNN), has only one input and one output layer. SLFFNN has no feedback connections; the input passed straight into output through various interconnections, which is regarded as a simple type of FFNNs. A representation of single-layer perceptron or SLFFNN presents in Fig. 15.

The sum products of SLFFNN weights and the inputs are computed in the node. An example of SLFFNN is a perception. In operation, perceptron usually returns a function based on inputs.

T. MULTILAYER NEURAL NETWORK

Multilayer neural networks (MLNNs) are a class of feedforward neural networks (FFNNs) that are performing great in PR. Multilayer (MLFFNNs) has only one input, one or more hidden layers, and one output layer without feedback relations example is MLP [201], [202]. MLFFNNs means multilayer perceptions (MLP), which are most studied and applied NN in practice. MLFFNNs solve the barrier of SLFFNN as they are capable of handling non-linear learning tasks. MLFFNN is represented in Fig. 16.

Fig. 16 demonstrates an architecture of an MFFNN of 3-2-1 network, 3 inputs, with 2 hidden layers, 1 output. Node output in ANN computation is mathematically expressed as;

$$y_j = \sum_{i=1}^{n} w_{ij} x_i + \theta_j \tag{2}$$

y_i represent parameter moved to subsequent layer let say node, j, then "n" termed amount of moving edges to node j, x_i xi's stand for those enter items through the unit to the node, j, θj equals bias node, j.

In the year 2001, for instance, Khadir, and Ringwood [203] investigated the possibility of applying feedforward ANNs, to PR for predicting pasteurization plant control. The result demonstrates that a pasteurization plant provides choice for prediction, optimization, and internal modeling complexity. A presentation of online MLFFNNs PR can be described in Fig. 17.

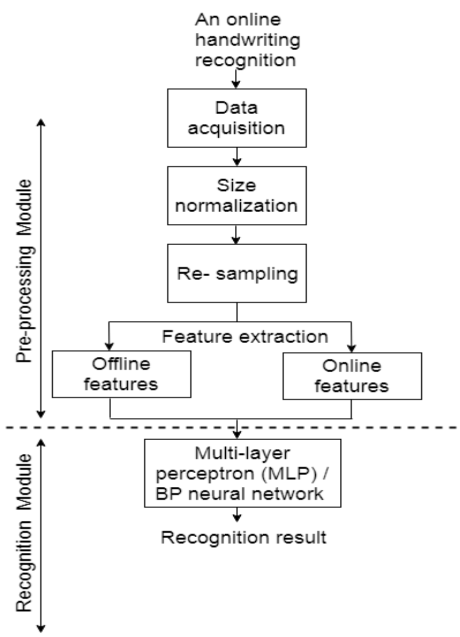


FIGURE 17. A multilayer neural network (MLNN) in PR.

ANN topology is explained in terms of the number of layers and nodes that features in layers. If ANN topology was ascertained, then the weight is allocated to each connection or edge for ANN to accurately identify unspecified, unknown, or hidden data during the learning period. The task of assigning or allocating weight into each is by learning in MLNBP.

Firstly, in a BPNN learning algorithm, biases and weights are allocated randomly in the range [0, 1] before utilized. After that, the network gives output according to the current condition of weights, just like in synaptic weights inhuman. Output compared to known good outputs, then mean squared error (MSE) determined. The cycle repeated until generalize error decreases to a known threshold. Subsequently, the network will continue to learn the task well. MFNNs method is ideal for handling demanding tasks in PR because of its high adaptability in a nonlinear system [204]–[207].

The difference between Single-layer feedforward neural networks (SLFFNNs), Multilayer feedforward neural networks (MLFFNNs) are the rest on neural network models enumerated in Section V are as follow;

SLFFNNs and MLFFNNs are the foundation of most neural network learning models. The neural network models like CNNs, RNNs, GANs, SAEs, DBN, RBMs, TDNNs, Reservoir computing, Transformer model, are just some special cases of feedforward networks and feedback networks that are applying to address a specific problem. These special cases of networks are mostly used for supervised machine learning tasks where the target function is already known. That is, the expected result from the network to be achieved, which are extremely important for practicing ML and form the basis of many commercial application areas. That is, application areas such as computer vision and natural language processing were highly affected or influenced by the presence of these NN models.

The difference characteristics of these NNs indicate that they can enhance each other to provide solutions to the problem that none of the intelligent systems alone can achieve with less system complexity. Therefore, because they can enhance each other and provide a solution to a complex task, combining or integrating two or more of these models become important for researchers and practitioners to achieve an objective.

U. A GENERAL REGRESSION NEURAL NETWORK

Another kind of ANN for PR is general regression NN (GRNN), which is a variation to RBF networks. GRNN is a memory-based network that enables the evaluation of increasing variables that converge to underlying nonlinear or linear regression facet. The GRNN is a single pass learning paradigm with a high parallel structure or a profoundly parallel structure that is taking from the input side to the output side [208]. GRNN functions well on disorderly representation than back-generation. NNs enhance performance if hidden layers increase to a certain level. Neurons require to be large enough to present the problem domain adequately.

Similarly, neurons are necessary to be small to enable generalization from trained data. However, a balance should be maintained between the network size and network sizing complexity. Remarkably, NN percentage of image recognition accuracy can be enhanced, using tansig activation functions (TAF) for a hidden layer (HL) and output layer (OL) than other approaches [214, 215].

VI. A FEEDFORWARD NEURAL NETWORK

The reason for the network termed feedforward is that, information flow in the forward direction. As x, is used to calculate some intermediate function in the hidden layer and then this calculated function in the hidden layer is used to calculate y. In this process, if the feedback is added from the last hidden layer to the first hidden layer it then represents RNN, as shown in Fig. 11.

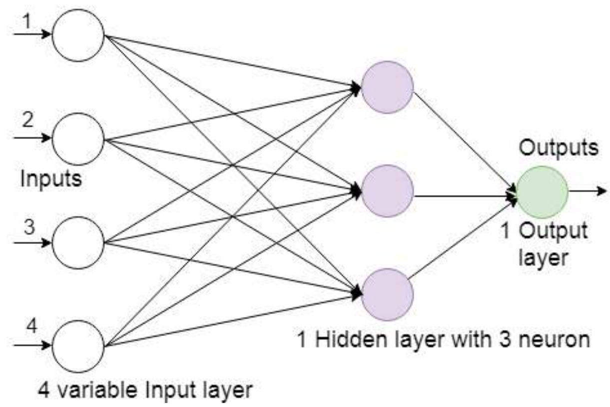


FIGURE 18. Algorithmic design of FFNN.

The goal of a feedforward network is to approximate some function f^*. For example, a regression function $y = f^*(x)$ maps an input x, to a value y. Then a feedforward network defines a mapping $y = f(x;\theta)$ and learns the value of the parameters "θ" which is the result of the best function approximation.

Feedforward neural networks (FFNNs) are also called Multilayered Network of Neurons (MLN). The FFNN was the first and simplest type of ANN models. In FFNN, the information moves from the input layer to any hidden layers and passes to the output layer without cycles/loops. FFNNs can be constructed with different types of units, like the binary McCulloch-Pitts neurons, the perceptron, continuous neurons frequently with sigmoidal activation are used in the context of backpropagation.

Most ANNs, like FFNNs, has no memory of the input that they received in a short while ago. A common family of ANNs for pattern classification (PC) is FFNNs like MLP and RBF [211]. Three elements influence the success of the PR techniques; the volume of data, the method applied, the designer, and the user.

The task in PR is building systems with the capability to handle large data. In FFNNs like multilayer perceptron (MLP) with a two-layer pattern, when the function of output is stepped, before implementing classification task; however, if the output function becomes linear, it does regression task. Different types of NNs are used for PC, but it depends on the application requirement. Feed-forward backpropagation NN (FFBPNN) is exploited to execute non-linear differentiable functions. Increasing the learning rate in FFBPNN leads to a decreasing in convergence time [212]. FFBPNN does not work efficiently where access to information is broad. A typical design sample of FFNN is shown in Fig. 18.

Fig. 18 represents the design of FFNN with a circle in the network that constitutes neurons in the ANN. ANNs permit information to be transmitted only in one direction from input to output. During information transmitting there is no looping (feedback), the output of a layer does not disturb other layers. However, information transmitting tends to be one-directional and straightforward in the networks that interrelate inputs and outputs.

In FFNN, all data flow the same direction, i.e. from left to right. Also, the bottom layer of inputs is not taken as an actual FFNN layer. The movement of information is to the right from the left-hand side, as shown in Fig. 18. Input, hidden, or output variables represent nodes, graph edges equivalent to adaptive parameters [208]. FFNN analytic function could be written in mathematical term to show the output of jth hidden node in linear combination (LC) weight of "n" with input values "x_i" as follow;

$$a_j = \sum_{i=1}^{n} ujixi + bj \tag{3}$$

Perceptron models are like "logic gates" that can recognize and discriminate to achieve a specific function. A perceptron sends a signal or not send a signal concerning weighted inputs. Another SLNN is a "single-layer binary linear classifier" that isolate or seclude inputs into one group or two groups. Hidden parameter "j" obtained by converting linear summation of (3) with an AF $g(\cdot)$ to give;

$$z_j = g(a_j) \tag{4}$$

Also, the LCs of the hidden parameters generate network outputs in form;

$$a_k = \sum_{j=1}^{m} vkjzj + ck \tag{5}$$

Variables $\{u_{ji}, v_{kj}\}$ represent *weights*, variables $\{b_{j,} c_k\}$ termed *biases*. Biases and weights represent adaptive variables in the network. One to one relationship exists between parameters like those in analytic function, nodes, and edges, respectively in the graph. Feedback NNs (FBNNs) are dynamic; having changes of state constantly until stability.

Also, the network remains stable until input changes, then a new stable point located. More also, feedback designs are called recurrent or interactive, recurrent referred to feedback interconnections in single-layer organizations like single-layer perceptron.

FFNN is utilized in spaces where classic machine learning (ML) methods applied; the main achievement of FFNN has been in speech recognition, visual imagery, and computer vision. One popular real-world application can be found in self-driving cars and Google. An application challenge of an FFNN is a single layer of any function with a larger single layer may fail to learn or generalize correctly.

Currently, popular FFNN applications to PR include; single-layer FFNNs, multiple layer FFNNs or multi-layer and recurrent NNs, Probabilistic, Convolutional, General regression NN, SAE, GAN, and TDNN. Other feedforward NN include; a deep stacking network (DSN), Tensor deep stacking networks.

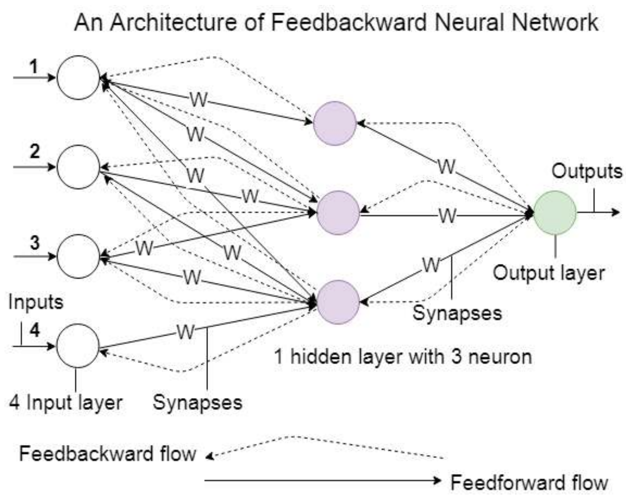


FIGURE 19. Algorithmic design of FBNN.

VII. A FEEDBACKWARD NEURAL NETWORK

ANNs permit information to be transmitted in two directions, that is, from the input to the output and back to the input. During information transmitting, there is looping (feedback). Hence this kind of network is termed backpropagation or feed backward neural network (FBNN). An ANN can learn from the wrongdoing, and right doing this is called feedback. Feedback is how human being learns from what is wrong and right. Similarly, this is what an ANN needs to learn.

Neural networks (NNs) learn things in exactly the way as the human brain, and typically by a feedback process called backpropagation (shortened as "backdrop"). The output of the network can be compared with the output it was initially meant to produce. using the differences between the outputs to modify the weights of the network connections between the neurons. Next is to work from the output units through the hidden neurons back to the input neurons by going backward. The family of FBNNs like Kohnen network is utilized for feature mapping and data clustering. Now from the gained knowledge of FFNN structure in Fig. 18, one can easily represent FBNN. A typical architecture of FBNN is shown in Fig. 19.

Backpropagation referred backward propagation of errors, which is an algorithm for supervised learning of ANNs using gradient descent. That is a first-order iterative optimization algorithm for finding the minimum of a function. Given an ANN and an error function, the algorithm computes or calculates the gradient of the error function with respect to the NN's weights (Ws).

In an algorithmic design functionality of FBNN, the information flow forward to learn and then backward to retain information learned as illustrated in Fig. 19. Thus, over time, backpropagation causes the network to learn effectively by bridging the gap between the output and the intended output.

TABLE 1. Comparison of ANN models for pattern recognition in percentage.

ANN models analysis on application to PR from 500 articles.	Number of articles (%)
Recurrent neural networks (RNNs)	29.38%
Convolutional neural networks (CNN)	10.21%
Multilayer perceptron	8.0%
Single-layer perceptron	7.7%
Hopfield network	3.0%
Hybrid network model	2.6%
Deep direct reinforcement learning	2.3%
Sparse auto-encoder (SAE)	2.1%
Competitive network model	2.1%
Time-delay neural network (TDNN)	2.0%
Deep belief nets (DBNs)	1.9%
Generative adversarial networks (GANs)	0.6%
Probabilistic neural network (PNN)	0.5%
Transformer models	0.1%
Self-organizing map (SOM)	5.26%
Radial basis function networks (RBFNs)	6.95%
Restricted Boltzmann machine (RBM)	3.10
Deep recursive network application	8.0%
Machine learning neural network (MLNN)	2.1%
Reservoir computing	2.1%
Total	**100%**

TABLE 2. Comparative experimental result on ANN models for pattern recognition.

Performance ANN models analysis on application to PR	Accuracy	Accuracy (%)
Recurrent neural networks (RNNs)	0.8389	83.39
Convolutional neural networks (CNN)	0.8376	83.76
Multilayer perceptron	0.8356	83.56
Single-layer perceptron	0.8254	82.54
Hopfield network	0.8126	81.26
Hybrid network model	0.8976	89.76
Deep direct reinforcement learning	0.8879	88.79
Sparse auto-encoder (SAE)	0.8978	89.78
Competitive network model	0.8976	89.76
Time-delay neural network (TDNN)	0.7976	79.76
Deep belief nets (DBNs)	0.8978	89.78
Generative adversarial networks (GANs)	0.8996	89.96
Probabilistic neural network (PNN)	0.7576	75.76
Transformer models	0.8997	89.97
Self-organizing map (SOM)	0.8175	81.75
Radial basis function networks (RBFNs)	0.8353	83.53
Restricted Boltzmann machine (RBM)	0.8040	80.40
Deep recursive network application	0.8997	89.97
Machine learning neural network (MLNN)	0.7011	70.11
Reservoir computing	0.8353	83.53

That is, making the gap between the outputs smaller to the point where the two exactly matches, therefore, the neural network learns the precise or correct output.

A studied have shown that over 80% of NN projects use BP to solve diverse challenges [160], [213]–[215]. MLANN with learning by BP algorithm is mostly used to solve rain attenuation obstacles due to its success in training and remarkable performance in PR [216]. Examples of feed backward NN that is widely used in the research area of PR are; RBM, SOM, Hopfield, RNN, Reservoir computing, and Long short-term memory (LSTM).

VIII. RESULTS

In evaluating the performance of ANN models, some statistical indicators for measuring the performance of ANN in many studies was adopted. These statistical indicators were used as evaluation metrics and validation functions to understand success in ANN. The indicators include absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE), a variance of absolute percentage error (VAPE) and R-squared.

Percentage of the Number of Papers Studied: The number of articles studied was 500 articles on ANN models application to pattern recognition as compared in Table 1.

Table 1 shows that Recurrent neural networks (RNNs), Convolutional neural networks (CNN), Multilayer perceptron, and Single-layer perceptron are the most used ANN models for pattern recognition. Though, that does not imply that they are the best because there are other improved models that may perform better.

Most of the new models that were discovered to address the limitations or problem of the former models are yet to be widely harnessed like GAN, SAE, DBN, and Transformer. However, a deep recursive network model seems to be popular in application compared to either GAN, SAE, DBN, PNN, or Transformer model.

Comparative of Experimental Result on ANN Models: The experimental result on accuracy are compared and presented. Similarly, the experimental result on the percentage of accuracy is compared and presented. Therefore, for each ANN model studied the result were compared using a distinct dataset from several articles studied on their application to PR and are represented in Table 2.

Table 2 shows the various levels of accuracy on the performance of ANN models application to PR. Although, each ANN model level of accuracy varies with the kind of applied model. Each model's performance depends on the specific problem addressing. However, each model can be tested with

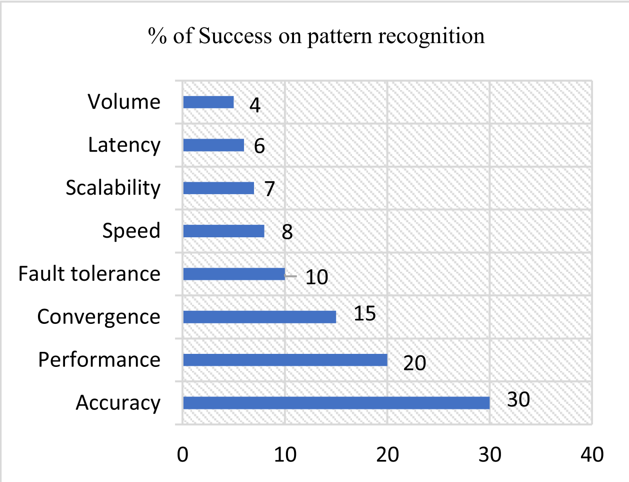


FIGURE 20. Articles classification on various data analysis influence on PR.

the same data for the same problem to know the model that can perform better than the other. Importantly, each model can be effective in its various applications in practice.

The Result on Data Analysis Influence: Distribution of articles over choosing topics on data analysis influence with respect to data volume, processing speed, performance, processing, scalability, latency, convergence, and fault tolerance was evaluated and presented as the result of data analysis influence for ANN application to PR. The evaluation shows that there are about 4/8ths of the investigations that addressed accuracy problems. About 3/8ths of the articles focused on performance issues, and approximately 1/8ths addresses convergence challenges in ANN application to PR. That is, about 15% of current studies focus on convergence challenges in ANN application to PR. While processing speed scalability, and latency were 8%, 7%, and 6% respectively, meanwhile, the volume of data scored 4%. The research samples containing 500 articles, as displayed in Fig. 20.

Fig. 20 result was based on the evaluation and validation functions like mean absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE) and variance of absolute percentage error (VAPE). The fields of ANN applications to PR are representing in Fig. 21.

As showed in Fig. 21, thus, the ANN reveals successes in recent times in diverse fields of application [217] to security, science, medical science, engineering, agriculture, finance, video and images, manufacturing, energy, arts, business, and management, etc. The wide area of ANN usefulness to humanity and many fields not listed and not yet discovered in the area of ANN application to PR. The interconnection among distinct subjects demonstrate that ANN can be applying to many industries, disciplines, and professions. Hence, interested individuals and a group of persons can explore many gaps and areas of studies for a better solution and more successes.

Summary of Result on the Areas of ANN Applied to PR: The result from areas of ANN applied to PR can be summarized.

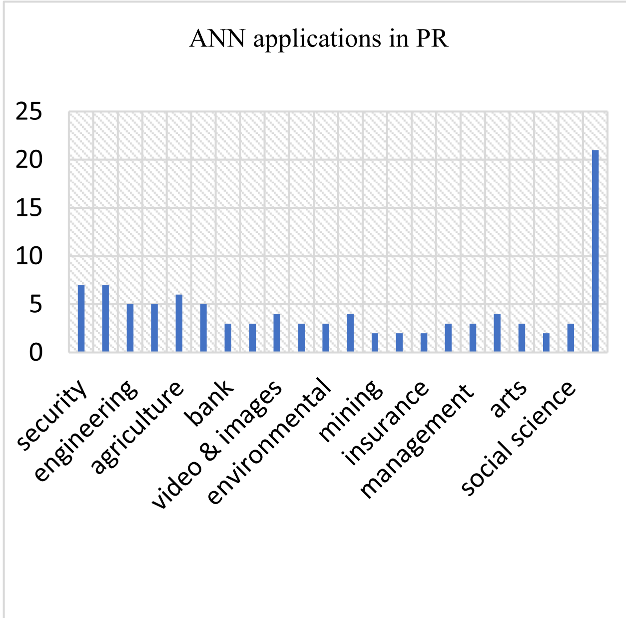


FIGURE 21. The framework of ANN application Pattern.

TABLE 3. On ANN application to pattern recognition.

Field of ANN application	% score on ANN PR
Security	7
Science	7
Engineering	5
Medical science	5
Agriculture	6
Finance	5
Bank	3
Weather and Climate	3
Video & Images	4
Education	3
Environmental	3
Energy	4
Mining	2
Policy	2
Insurance	2
Marketing	3
Management	3
Manufacturing	4
Arts	3
Transportation	2
Social science	3
Other fields of application	21

The ANN applied to PR highlight its current usefulness in the various fields, as shown in Table 3.

Other fields of application in Table 3 indicate that ANN is applicable in many fields apart from the few ones listed. That is, there are many areas of ANN application that are yet to explore or discover, that means, there are prospects in the use of ANN, especially in its application to PR. ANNs application to PR in many fields and disciplines of having proven success such as medical science, computer vision, engineering, fishing, agriculture, graphics, arts, social science, manufacturing, business, management, security, telecommunication, etc.

IX. DISCUSSION

Artificial neural network concepts and rules have great potential for resolving problems, as demonstrated in many practices. FFNN propagation has more potential for PR, but to optimize BP NNs performance, there must be systematic adoption in the model construction process. PR can be different depending on the kind of patterns that can be recognized, the kind of data, and the amount of data one is expecting, etc. Current and previous studies of PR with other techniques can be compared to ANN models, like the used of structural, statistical, template matching, fuzzy, and hybrid techniques. However, ANN application to PR has significant advantages over many other methods as previously mentioned in the introduction, like been a good recognizer of patterns and been a robust classifier.

An advantage of ANNs to statistical modeling is that it does not require rigidity structured designs and map function with Incomplete data. Notwithstanding, many issues in PR through ANN needs addressing. Recently, DNNs attracted research attention specifically in image recognition. During PR, contrary to the statistical model, ANN models rely much on the quantity of training data, presentation of samples, and procedures for training. In achieving a good recognition of pattern, a good feature must be selected with enough multi-classifier systems. ANN computational intelligence help in addressing different problems and minimizes hazard. Though, learned features employed in DNN to achieved superb success in PR than hand-crafted methods.

Notwithstanding, recent development showed that PR could be better achieved using carbon distribution by carbon's bonding patterns, uniquely, amino acids efficient to conserve carbon distribution. Although contemporary protocol on carbon distribution is yet to be discovered, present-day, carbon can be determining factor information of PR [218]–[222]. Researchers can explore using carbon for PR as an alternative.

Result of Comparison With Other State-of-the-Art Studies: Comparison of the study with other state-of-the-arts studies is presented in Table 4.

X. AREAS IMPROVEMENT IN ANN APPLICATION TO PR

Meanwhile, other noticeable areas of improvement in ANN PR are as follow;

1. Normalization of data is needed in scaling to a range commensurable with a transfer function in the output layer.

2. In knowing actual model input. The inputted parameter can be calculated with a preliminary knowledge of computational approaches like stepwise model-building and cross-correlation analysis.

3. An HL should be sufficient for a decision on network geometry since it is appropriate in practice. However, mathematical connections can use in locating an upper level of HL nodes necessary to make ANN equivalent a relationship that increases. Excluding non-convergence methods, for example, cross-validation, an association among training samples and HL nodes required.

4. In improving PR in ANN, it is necessary to know the character of a network at the parameter specification and training stage because it will help to count a local minimum of the error surface.

5. Validation of ANN models need improvement in PR; this is important for practical scenarios and standard in ANNs applications.

XI. RESEARCH OPPORTUNITIES AND FUTURE DIRECTIONS

(i) Research attention requiring to discover more wide contributions of ANN's PR problems like recognizing complex patterns that have an arbitrary orientation.

(ii) Further research is requiring in the ANN object classification problem for a better result.

(iii) More study is needed to tackle an object location problem when applying ANN models. Particularly, a study to enhance the locational object problem.

(iv) Extensive research is requiring fixing scaling or normalization issues when using ANN models. An example is running the ANN program without normalization.

Though scaling or (normalization) may not be operational necessary for NNs training, however, scaling helps in transposing input parameters to data. Some people argued that normalization is not required for properly coded ANNs. Hence, this argument required more investigation for clarification or a better result.

(v) Researchers can apply intelligent analysis to improve the performance of ANN in its application to PR. Like, applying backpropagation NNs, probabilistic NN, supervised associating networks, MLPNN architectures, learning vector quantization, MLNN, and hybrid NN models.

(vi) Further research is requiring on the issues with ANN-based control chart PR to achieve better performance, since control chart PR has become an active area in recent times [223]–[227].

(vii) In nowadays, global computing, social media connectivity, most web users are passionate to share and engage in an online synergetic data analysis [226]. Therefore, there is a prospect in the use of ANN for cognitive cloud computing, especially in the aspect of data analysis and PR.

(viii) ANN capability to learn by example makes them great, powerful, and flexible. Researchers can apply AI mechanisms to accomplish immense success in many areas of endeavors, including the manufacturing industry [228]. Particularly, the study focus is requiring in the area of fault recognition in machines, fault recognition in plants including power and energy like the nuclear power plant, wind, solar, hydroelectricity, coal, gas, etc. ANN fault recognition in quality recognition, and PR in products.

Likewise, ANN widely applied in other industries like in real estate management, security, in business.

TABLE 4. Comparison of survey with some state-of-the-arts surveys.

Number of survey papers	Comprehensive Analysis of ANN application to pattern recognition	Pattern recognition case study	Mapping of pattern recognition vectors	Applied single methods	Applied combine approaches	Identified new pattern recognition techniques	Recommended approaches	Challenges	Research opportunities
[10]	-	✓	✓	✓	-	✓	-	✓	-
[12]	✓	✓	-	✓	✓	✓	✓	✓	✓
[36]	-	-	✓	-	✓	✓	✓		-
[102]	✓	-	✓	✓	-	-	-	✓	✓
[120]	✓	✓	-	✓	-	-	-	✓	-
New Survey	✓	✓	✓	✓	✓	✓	✓	✓	✓

The study focus is required in the area of market price PR in machines on housing in the estate industry, then purchase of material, and product in security and business industry respectively for quick decision making.

(ix) Research collaborations are requiring manufacturing industries and domain scientists. That is scientists like computer scientists, IT experts, engineers, and, statisticians for better performance and success of ANN application to pattern recognition in industrial development.

(x) Recently, like any field in science, astronomy is experiencing an exponential increase in data volume, complexity, and quality. Therefore, research focus is requiring in Astronomy data-intensive challenges for a better result.

(xi) The research is requiring in an adversarial attack on machine misinterpretation of PR.

(xii) The future research will enable a solution to analysis problems in advance digital sky exploit in the context of the virtual observatory. An example includes issues in the operating star-galaxy category in heterogeneous extensive visual representation datasets.

(xiii) There is a need for study into improving the method of selecting parameters for ANN design concerning PR.

(xiv) Future research should be focused on the problem area in ANN application to PR. This includes challenges relating to the real-world issues on pattern computational. Though it may need to kneel knowledge of information processing from the biological and psychological perspective.

(xv) It is crucial to design new models of artificial neurons as processing paradigm, their associations, activation, synaptic, learning rules, and memorable algorithms so that they can produce a remarkable success in many fields of human endeavor.

(xvi) New research expected in the building of fundamental structures that will solve a category of PR challenges that can lead to the formation of a foundation block for the advancement of new ANN architectures.

(xvii) The area of architecture requires a concept to develop new ANN architectures from the known principles, compositions, and structures to unknown principles that can solve complex and multi-complex pattern recognition challenges. The designs should take cognizant of the problems at hand based on suitable architectures that can address it.

(xviii) New research expecting in application area to address a given practical problem using the principles of ANN, but with knowledge from other areas such as signal processing, applied physics, geography, history, arts, economic, etc.

(ixx) Further research into, problems, issues, challenges, and concerning PR in computer vision and natural language research.

[xx] In recent times, advanced persistent threats (APTs), have been targeting nations, organizations and private sectors. The rate at which the APTs attack tools and techniques are evolving is making any existing security measures inadequate is worrisome. As defensive measures are put in place to secure all aspect computer system, also, the attackers are finding new strategies to penetrate their target computer systems.

In addressing the problem of APT attack there have been popular methods used for security of a networked system such as; Support Vector Machines (SVM), random forest model, Open Source version of Security Information and Event Management (SIEM), etc. But, some of these methods APT measures in practice have not been wholly successful. Therefore, more research is expecting on APTs to recognize they are the pattern in the network and to exclusively address the problem of long-time detection in a victim network system. Also, to prevent data damage and data exfiltration from the victim computer.

[xxi] In recent years, Malware become a large and well-organized market. This rapid growth of the Malware market is affecting the security of the computer industry. The increasing presence of IoT devises in a wide range of applications, processing, and computing capabilities without effective security protection make them vulnerable to Malware attack target. Also, the Internet of Things (IoT) devices is increasing for

different purposes such as remote sensing, collecting data in both military and civilian environments.

Malware attacks are causing data and financial losses to the users of computer systems which require anti-malware mechanisms for funding and deactivating invasion in the network system. Thus, more study is requiring into Malware PR in a computer system using ANN to solve the problem of early detection of an attack.

[xxii] Similarly, Botnet attacks are causing data and financial losses to the users of computer network systems that require urgent a solution such as using ANN to handle the challenge of early detection of an attack via PR.

[xxiii] More study is requiring in the forensic investigation on PR of a crime using ANN to overcome the problem of data collection and data analysis methodology on a crime case for evidence in court.

[xxiv] Study is needing in nuclear security using ANN to address the problem of fault recognition during operation in nuclear power plants against an accident.

XII. CONCLUSION

This review has summarized current and previous trends in PR, with a specific focus on performance. New development in the ANN model for PR was comprehensively discussed, such as the use of SAE, GAN, DBN, RBM, TDNN, Reservoir computing, Transformer model, etc. Likewise, current and previous development in the use of ANN models such as CNN, RNN, SLP, MLP, and SOM was highlighted.

Some studies empirically show that accurate information transformation of input data is important for satisfactory performance. Some other studies reveal that temporal and spatial scaling of input data critically affects computational performance. The study shows that for the past two decades, ANN applications to PR has achieved significant success and is becoming more popular and widen.

The review provides a better understanding of why ANN models are used in diverse applications to PR and at the same time, be useful to many fields in the nearest future. With the comprehensive study, summary, and excellent cognitive computational approach of ANN to PR, one can propose the best performing models for future applications that can address many challenges in PR. Using more than one ANN model, the recognizer can solve simple, complex, and multi-complex problems. Also, designing a new specific model for a particular problem can result in a better solution. Furthermore, problems, issues, challenges, and research directions highlighted can be useful in computer vision and natural language research.

This comprehensive review can be useful to current researchers as a starting point in facilitating further advancement in the field, especially in addressing the issue of ANN application to PR. Similarly, this work can be helpful to a newcomer in the field, as many topics and problem areas are available to explore. Meanwhile, ANN application to PR has a bright future in diverse areas and disciplines.

Study limitation includes a lack of access to data used in many studies for relative comparison of results in similar and dissimilar areas of investigation. But some statistical indicators for evaluation and validations used in many studies were utilized in this study to understand success in ANN. These statistical indicators include absolute percentage error (MAPE), mean absolute error (MAE), root mean squared error (RMSE), a variance of absolute percentage error (VAPE) and R-squared.

Experimental results showed that ANN is becoming most popular comparable to using structural, statistical, template matching, fuzzy, and hybrid techniques to addressed pattern recognition problems in many fields. Hence, one can propose that more research focus should be on FFNN and FBNN models. Importantly, more research focuses on the current hotspots ANN models like SAE, GAN, CNN, DBN, RNN, RBM, TDNN, Reservoir computing, SLP, MLP, Transformer model, etc.

Furthermore, we propose that more research focus on the new design model of ANN to resolve many highlighted problems of PR task for a better result.

COMPLIANCE WITH ETHICAL STANDARDS

CONFLICTS OF INTEREST

The authors stated that they have no competing interests.

HUMAN AND ANIMAL'S RIGHTS

No studies with human participants or animals performed by any of the authors.

INFORMED CONSENT

In this article, informed consent was not required as no human or animals were involved.

REFERENCES

[1] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer*, vol. 29, no. 3, pp. 31–44, 1996.

[2] J. K. Basu, D. Bhattacharyya, and T. H. Kim, "Use of artificial neural network in pattern recognition," *Int. J. Softw. Eng. Appl.*, vol. 4, no. 2, pp. 1–13, 2010.

[3] H. Y. Priyangga and D. Ruliandi, "Application of pattern recognition and classification using artificial neural network in geothermal operation," in *Proc. 43rd Workshop Geothermal Reservoir Eng.* Stanford, CA, USA: Stanford Univ., Feb. 2018.

[4] C. Shang, A. Palmer, J. Sun, K.-S. Chen, J. Lu, and J. Bi, "VIGAN: Missing view imputation with generative adversarial networks," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2017, pp. 766–775.

[5] R. Bala and D. D. Kumar, "Classification using ANN: A review," *Int. J. Comput. Intell. Res.*, vol. 13, no. 7, pp. 1811–1820, 2017.

[6] S. Knerr, L. Personnaz, and G. Dreyfus, "Handwritten digit recognition by neural networks with single-layer training," *IEEE Trans. Neural Netw.*, vol. 3, no. 6, pp. 962–968, Nov. 1992.

[7] M. Nikzad, K. Movagharnejad, and F. Talebnia, "Comparative study between neural network model and mathematical models for prediction of glucose concentration during enzymatic hydrolysis," *Int. J. Comput. Appl.*, vol. 56, no. 1, pp. 1–6, 2012.

[8] L. Lazli and M. Boukadoum, "Hidden neural network for complex pattern recognition: A comparison study with multi-neural network based approach," *Int. J. Life Sci. Med. Res.*, vol. 3, no. 6, pp. 234–245, 2013.

[9] J. Kepka, "The current approaches in pattern recognition," *Kybernetika*, vol. 30, no. 2, pp. 159–176, 1994.

[10] P. S. M. Kaur, "Classification in pattern recognition: A review," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 4, pp. 1–9, 2013.

[11] R. Amardeep and K. T. Swamy, "Training feed forward neural network with backpropogation algorithm," *Int. J. Eng. Comput. Sci.*, vol. 6, no. 1, 2017.

[12] A. Kattan and R. Abdullah, "Training of feed-forward neural networks for pattern-classification applications using music inspired algorithm," *Int. J. Comput. Sci. Inf. Secur.*, vol. 9, no. 11, p. 44, 2011.

[13] S. He, R. W. H. Lau, W. Liu, Z. Huang, and Q. Yang, "SuperCNN: A superpixelwise convolutional neural network for salient object detection," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 330–344, 2015.

[14] J. Ruiz-del-Solar, P. Loncomilla, and N. Soto, "A survey on deep learning methods for robot vision," 2018, *arXiv:1803.10862*. [Online]. Available: https://arxiv.org/abs/1803.10862

[15] S. Li and W. Deng, "Deep facial expression recognition: A survey," 2018, *arXiv:1804.08348*. [Online]. Available: https://arxiv.org/abs/1804.08348

[16] B. Hasani and M. H. Mahoor, "Facial expression recognition using enhanced deep 3D convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2017, pp. 30–40.

[17] P. Khorrami, T. Paine, and T. Huang, "Do deep neural networks learn facial action units when doing expression recognition?" in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Dec. 2015, pp. 19–27.

[18] S. A. Bargal, E. Barsoum, C. C. Ferrer, and C. Zhang, "Emotion recognition in the wild from videos using images," in *Proc. 18th ACM Int. Conf. Multimodal Interact.*, Oct. 2016, pp. 433–436.

[19] R. Vazquez, "Izhikevich neuron model and its application in pattern recognition," *Austral. J. Intell. Inf. Process. Syst.*, vol. 11, no. 1, pp. 35–40, 2010.

[20] R. Hühnerbein, F. Savarino, F. Åström, and C. Schnörr, "Image labeling based on graphical models using wasserstein messages and geometric assignment," *SIAM J. Imag. Sci.*, vol. 11, no. 2, pp. 1317–1362, 2018.

[21] X. Zeng, W. Ouyang, J. Yan, H. Li, T. Xiao, K. Wang, Y. Liu, Y. Zhou, B. Yang, Z. Wang, H. Zhou, and X. Wang, "Crafting GBD-net for object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 9, pp. 2109–2123, Aug. 2018.

[22] P. Sala and S. Dickinson, "3-D volumetric shape abstraction from a single 2-D image," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Dec. 2015, pp. 1–9.

[23] K. Grauman, "Efficiently searching for similar images," *Commun. ACM*, vol. 53, no. 6, pp. 84–94, 2010.

[24] S. S. Dahikar and S. V. Rode, "Agricultural crop yield prediction using artificial neural network approach," *Int. J. Innov. Res. Elect., Electron., Instrum. Control Eng.*, vol. 2, no. 1, pp. 683–686, 2014.

[25] K. Jha, A. Doshi, P. Patel, and M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," *Artif. Intell. Agricult.*, vol. 2, pp. 1–12, Jun. 2019.

[26] G. Niedbała, "Application of artificial neural networks for multi-criteria yield prediction of winter rapeseed," *Sustainability*, vol. 11, no. 2, p. 533, 2019.

[27] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[28] A. Seema and D. Rajeshwar, "Pattern recognition techniques: A review," *Int. J. Comput. Sci. Telecommun.*, vol. 3, no. 8, Aug. 2012.

[29] L. Siegel, "A procedure for using pattern classification techniques to obtain a voiced/Unvoiced classifier," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 27, no. 1, pp. 83–89, Feb. 1979.

[30] K.-S. Fu, "Pattern recognition and image processing," *IEEE Trans. Comput.*, vol. C-100, no. 12, pp. 1336–1346, Dec. 1976.

[31] L. Zheng and X. He, "Classification techniques in pattern recognition," Dept. IT, Univ. Technol., Sydney, NSW, Australia, Tech. Rep., 2005.

[32] K. A. Qazi, T. Nawaz, Z. Mehmood, M. Rashid, and H. A. Habib, "A hybrid technique for speech segregation and classification using a sophisticated deep neural network," *PloS ONE*, vol. 13, no. 3, 2018, Art. no. e0194151.

[33] R. Mehta and T. Arbel, "RS-Net: Regression-segmentation 3D CNN for synthesis of full resolution missing brain MRI in the presence of tumours," 2018, *arXiv:1807.10972*. [Online]. Available: https://arxiv.org/abs/1807.10972

[34] D. B. Hanchate, M. Nalawade, M. Pawar, V. Pophale, and P. K. Maurya, "Vocal digit recognition using artificial neural network," in *Proc. 2nd Int. Conf. Comput. Eng. Technol. (ICCET)*, vol. 6, Apr. 2010, pp. V68–V88.

[35] D. Schuurmans and M. A. Zinkevich, "Deep learning games," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2016, pp. 1678–1686.

[36] G. Wei, L. Piyan, Z. Hai, and M. Zhan, "The research and application of image recognition based on improved BP algorithm," in *Proc. 3rd Int. Conf. Intell. Netw. Intell. Syst. (ICINIS)*, Nov. 2010, pp. 80–83.

[37] R. D. Dony and S. Haykin, "Neural network approaches to image compression," *Proc. IEEE*, vol. 83, no. 2, pp. 288–303, Feb. 1995.

[38] W. Chang, W. Chung, H. Chen, and S. Chou, "An international perspective on fighting cybercrime," in *Proc. Int. Conf. Intell. Secur. Inform.* Berlin, Germany: Springer, Jun. 2003, pp. 379–384.

[39] N. Wang, D. Tao, X. Gao, X. Li, and J. Li, "A comprehensive survey to face hallucination," *Int. J. Comput. Vis.*, vol. 106, no. 1, pp. 9–30, 2014.

[40] Y. LeCun and Y. Bengio, "Convolutional networks for images, speech, and time series," in *The Handbook of Brain Theory and Neural Networks*, vol. 3361, no. 10. Cambridge, MA, USA: MIT Press, 1995.

[41] T. V. Nguyen, Q. Zhao, and S. Yan, "Attentive systems: A survey," *Int. J. Comput. Vis.*, vol. 126, no. 1, pp. 86–110, 2018.

[42] Y. Wu and Q. Ji, "Facial landmark detection: A literature survey," *Int. J. Comput. Vis.*, vol. 127, no. 2, pp. 115–142, 2017.

[43] A. A. Khodaskar and S. A. Ladhake, "Pattern recognition: Advanced development, techniques and application for image retrieval," in *Proc. Int. Conf. Commun. Netw. Technol. (ICCNT)*, Dec. 2014, pp. 74–78.

[44] K. Nakajima, H. Hauser, T. Li, and R. Pfeifer, "Exploiting the dynamics of soft materials for machine learning," *Soft Robot.*, vol. 5, no. 3, pp. 339–347, 2018.

[45] A. Globerson, E. Stark, E. Vaadia, and N. Tishby, "The minimum information principle and its application to neural code analysis," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 9, pp. 3490–3495, 2009.

[46] A. Abran, J. W. Moore, P. Bourque, R. Dupuis, and L. L. Tripp, "Guide to the software engineering body of knowledge," to be published.

[47] B. R. Gaines and P. V. Facey, "Some experience in interactive system development and application," *Proc. IEEE*, vol. 63, no. 6, pp. 894–911, Jun. 1975.

[48] S. O. Markel, "System and method for distribution of interactive content to multiple targeted presentation platforms," U.S. Patent 7 162 697, Jan. 9, 2007.

[49] E. Gelenbe, "Random neural networks with negative and positive signals and product form solution," *Neural Comput.*, vol. 1, no. 4, pp. 502–510, 1989.

[50] K. Gurney, *An Introduction to Neural Networks*. Boca Raton, FL, USA: CRC Press, 2014.

[51] W. Shao, H. Luo, F. Zhao, Y. Ma, Z. Zhao, and A. Crivello, "Indoor positioning based on fingerprint-image and deep learning," *IEEE Access*, vol. 6, pp. 74699–74712, 2018.

[52] S. T. Saste and S. M. Jagdale, "Comparative study of different techniques in speaker recognition," *Int. J. Adv. Eng., Manage. Sci.*, vol. 3, no. 3, pp. 1–4, 2017.

[53] J. Kwon and K. M. Lee, "A unified framework for event summarization and rare event detection from multiple views," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1737–1750, Sep. 2015.

[54] G. Paragliola and A. Coronato, "Gait anomaly detection of subjects with Parkinson's disease using a deep time series-based approach," *IEEE Access*, vol. 6, pp. 73280–73292, 2018.

[55] T. H. Kim, "Pattern recognition using artificial neural network: A review," in *Proc. Int. Conf. Inf. Secur. Assurance*. Berlin, Germany: Springer, 2010, pp. 138–148.

[56] P. Sao, R. Hegadi, and S. Karmakar, "A literature review on approaches of ECG pattern recognition," *Int. J. Inf. Sci. Intell. Syst.*, vol. 3, no. 2, pp. 79–90, 2014.

[57] A. K. Jain, R. P. W. Duin, and J. C. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[58] M. Imran, "Per layer time computation of CNN training with tensorflow," M.S. thesis, Dept. Comput. Sci. Eng., Politecnico Milano, Milan, Italy, 2018.

[59] A. Vieira and B. Ribeiro, *Introduction to Deep Learning Business Applications for Developers*. New York, New York, USA: Apress, 2018.

[60] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: An overview and application in radiology," *Insights Imag.*, vol. 9, no. 4, pp. 611–629, 2018.

[61] C. Yao, X. Bai, W. Liu, Y. Ma, and Z. Tu, "Detecting texts of arbitrary orientations in natural images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1083–1090.

[62] K. Saravanan and S. Sasithra, "Review on classification based on artificial neural networks," *Int. J. Ambient Syst. Appl.*, vol. 2, no. 4, pp. 11–18, 2014.

[63] M. Bouzenada, M. C. Batouche, and Z. Telli, "Neural network for object tracking," *Inf. Technol. J.*, vol. 6, no. 4, pp. 526–533, 2007.
[64] L. He, X. Ren, Q. Gao, X. Zhao, B. Yao, and Y. Chao, "The connected-component labeling problem: A review of state-of-the-art algorithms," *Pattern Recognit.*, vol. 70, pp. 25–43, Oct. 2017.
[65] G. Montavon, S. Lapuschkin, A. Binder, W. Samek, and K. R. Müller, "Explaining nonlinear classification decisions with deep Taylor decomposition," *Pattern Recognit.*, vol. 65, pp. 211–222, May 2017.
[66] G. N. Yannakakis and J. Togelius, *Artificial Intelligence and Games*. New York, NY, USA: Springer, 2018.
[67] S. Fan, T. T. Ng, B. L. Koenig, J. S. Herberg, M. Jiang, Z. Shen, and Q. Zhao, "Image visual realism: From human perception to machine computation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 9, pp. 2180–2193, Sep. 2018.
[68] U. Asif, M. Bennamoun, and F. Sohel, "A multi-modal, discriminative and spatially invariant CNN for RGB-D object labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 9, pp. 2051–2065, Sep. 2018.
[69] W. Vaz, R. G. Landers, and U. O. Koylu, "Neural network strategy for driving behaviour and driving cycle classification," *Int. J. Electr. Hybrid Vehicles*, vol. 6, no. 3, pp. 255–275, 2014.
[70] P. T. Saito, R. Y. Nakamura, W. P. Amorim, J. P. Papa, P. J. de Rezende, and A. X. Falcão, "Choosing the most effective pattern classification model under learning-time constraint," *PLoS ONE*, vol. 10, no. 6, 2015, Art. no. e0129947.
[71] Z. Boger, "Implementing artificial neural networks in nuclear power plants diagnostic systems: Issues and challenges," in *Proc. IAEA Tech. Committee Meeting Diagnostic Syst. Nucl. Power Plants*, Istanbul, Turkey, 1998, pp. 149–160.
[72] S. S. Roy, M. Ahmed, and M. A. H. Akhand, "Noisy image classification using hybrid deep learning methods," *J. ICT*, vol. 18, no. 2, pp. 233–269, 2018.
[73] M. Salerno, G. Costantini, and M. Carota, "Neural network approach to problems of static/dynamic classification," Ph.D. dissertation, Dept. Eng., Univ. Rome, Rome, Italy, 2008.
[74] M. E. Hodgson, "What size window for image classification? A cognitive perspective," *PE RS-Photogramm. Eng. Remote Sens.*, vol. 64, no. 8, pp. 797–807, 1998.
[75] O. Kwon, H. G. Kim, M. J. Ham, W. Kim, G.-H. Kim, J.-H. Cho, N. I. Kim, and K. Kim, "A deep neural network for classification of melt-pool images in metal additive manufacturing," *J. Intell. Manuf.*, to be published.
[76] J. P. Jesan, "The neural approach to pattern recognition," *ACM Digit. Library, Ubiquity*, Apr. 2004.
[77] J. S. Bridle, "Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition," *Neurocomputing*. Berlin, Germany: Springer, 1990, pp. 227–236.
[78] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," in *Emerging Artificial Intelligence Applications in Computer Engineering*, vol. 160. Amsterdam, The Netherlands: IOS Press, 2007, pp. 3–24.
[79] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
[80] D. Lee and J. Hu, "Local state-feedback stabilization of nonlinear fuzzy systems with measurement errors in the state," in *Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ)*, Jul. 2016, pp. 1140–1146.
[81] B. M. Chen, T. H. Lee, K. Peng, and V. Venkataramanan, "Composite nonlinear feedback control for linear systems with input saturation: Theory and an application," *IEEE Trans. Autom. Control*, vol. 48, no. 3, pp. 427–439, Mar. 2003.
[82] N. Vlahović and G. Kvaščev, "Noise reduction by using autoassociative neural networks," in *Proc. 13th Symp. Neural Netw. Appl. (NEUREL)*, Nov. 2016, pp. 1–5.
[83] B. J. Fernandes, G. D. Cavalcanti, and T. I. Ren, "Constructive autoassociative neural network for facial recognition," *PLoS ONE*, vol. 9, no. 12, 2014, Art. no. e115967.
[84] S. Li, J. He, Y. Li, and M. U. Rafique, "Distributed recurrent neural networks for cooperative control of manipulators: A game-theoretic perspective," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 2, pp. 415–426, Feb. 2017.
[85] B. J. Fernandes, G. D. Cavalcanti, and T. I. Ren, "Autoassociative pyramidal neural network for one class pattern classification with implicit feature extraction," *Expert Syst. Appl.*, vol. 40, no. 18, pp. 7258–7266, 2013.
[86] M. Gori and A. Tesi, "On the problem of local minima in backpropagation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 1, pp. 76–86, Jan. 1992.
[87] S. Erfani and M. Malek, "Potential applications of neural networks in network management," in *Proc. 36th Midwest Symp. Circuits Syst.*, Aug. 1993, pp. 257–260.
[88] H. Li and S. Ding, "Research and development of granular neural networks," *Appl. Math. Inf. Sci.*, vol. 7, no. 3, pp. 1251–1261, 2013.
[89] D. Zheng, W.-F. Xie, X. Ren, and J. Na, "Identification and control for singularly perturbed systems using multitime-scale neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 2, pp. 321–333, Feb. 2017.
[90] M. Yousefnezhad and D. Zhang, "Decoding visual stimuli in human brain by using anatomical pattern analysis on fMRI images," in *Proc. Int. Conf. Brain Inspired Cogn. Syst.* Cham, Switzerland: Springer, Nov. 2016, pp. 47–57.
[91] B. K. Rajithkumar and H. S. Mohana, "Template matching method for recognition of stone inscripted Kannada characters of different time frames based on correlation analysis," *Int. J. Elect. Comput. Eng.*, vol. 4, no. 5, pp. 719–729, 2014.
[92] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, 2007.
[93] A. Shokoufandeh, Y. Keselman, M. F. Demirci, D. Macrini, and S. Dickinson, "Many-to-many feature matching in object recognition: A review of three approaches," *IET Comput. Vis.*, vol. 6, no. 6, pp. 500–513, Nov. 2012.
[94] X. Guo, X. Liang, and X. Li, "A stock pattern recognition algorithm based on neural networks," in *Proc. 3rd Int. Conf. Natural Comput. (ICNC)*, vol. 2, Aug. 2007, pp. 518–522.
[95] H. Sira-Ramirez, S. Ahmad, and M. Zribi, "Dynamical feedback control of robotic manipulators with joint flexibility," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 4, pp. 736–747, Jul. 1992.
[96] A. Tzes, P.-Y. Peng, and C.-C. Houng, "Neural network control for DC motor micromaneuvering," *IEEE Trans. Ind. Electron.*, vol. 42, no. 5, pp. 516–523, Oct. 1995.
[97] A. D. Kalafatis, L. Wang, and W. R. Cluett, "Identification of Wiener-type nonlinear systems in a noisy environment," *Int. J. Control*, vol. 66, no. 6, pp. 923–941, 1997.
[98] M. Malinowski, M. Rohrbach, and M. Fritz, "Ask your neurons: A deep learning approach to visual question answering," *Int. J. Comput. Vis.*, vol. 125, nos. 1–3, pp. 110–135, 2017.
[99] M. A. Anusuya and S. K. Katti, "Speech recognition by machine: A review," 2010, *arXiv:1001.2267*. [Online]. Available: https://arxiv.org/abs/1001.2267
[100] S. G. Koolagudi and K. S. Rao, "Emotion recognition from speech: A review," *Int. J. Speech Technol.*, vol. 15, no. 2, pp. 99–117, 2012.
[101] N. Washani and S. Sharma, "Speech recognition system: A review," *Int. J. Comput. Appl.*, vol. 115, no. 18, pp. 7–10, 2015.
[102] J. Kim, J. L. Kwon, and K. M. Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 1637–1645.
[103] H. Flynn, "What are some computer vision tasks that deep learning still does not tackle well?" Univ. Oxford, Oxford, U.K., Tech. Rep., Feb. 2015.
[104] H. Liu, J. Fang, X. Xu, and F. Sun, "Surface material recognition using active multi-modal extreme learning machine," *Cogn. Comput.*, vol. 10, no. 6, pp. 937–950, 2018.
[105] M. Anbar, R. Abdullah, B. N. Al-Tamimi, and A. Hussain, "A machine learning approach to detect router advertisement flooding attacks in next-generation IPv6 networks," *Cogn. Comput.*, vol. 10, no. 2, pp. 201–214, 2018.
[106] Y. Shen and J. Wang, "Robustness analysis of global exponential stability of recurrent neural networks in the presence of time delays and random disturbances," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 1, pp. 87–96, Jan. 2012.
[107] A. Bramson, A. Baland, and A. Iriki, "Measuring dynamical uncertainty with revealed dynamics Markov models," *Frontiers Appl. Math. Statist.*, vol. 5, p. 7, Feb. 2019.
[108] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, "Autoencoding beyond pixels using a learned similarity metric," 2015, *arXiv:1512.09300*. [Online]. Available: https://arxiv.org/abs/1512.09300
[109] S. Zhao, J. Song, and S. Ermon, "Towards deeper understanding of variational autoencoding models," 2017, *arXiv:1702.08658*. [Online]. Available: https://arxiv.org/abs/1702.08658

[110] Y. Ma, K. Liu, Z. Guan, X. Xu, X. Qian, and H. Bao, "Background augmentation generative adversarial networks (BAGANs): Effective data generation based on GAN-augmented 3D synthesizing," *Symmetry*, vol. 10, no. 12, p. 734, 2018.

[111] M. Li, O. Li, G. Liu, and C. Zhang, "Generative adversarial networks-based semi-supervised automatic modulation recognition for cognitive radio networks," *Sensors*, vol. 18, no. 11, p. 3913, 2018.

[112] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.

[113] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, and V. Lempitsky, "Domain-adversarial training of neural networks," *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 2030–2096, 2016.

[114] W. Fedus, M. Rosca, B. Lakshminarayanan, A. M. Dai, S. Mohamed, and I. Goodfellow, "Many paths to equilibrium: GANs do not need to decrease a divergence at every step," 2017, *arXiv:1710.08446*. [Online]. Available: https://arxiv.org/abs/1710.08446

[115] F. Tramèr, A. Kurakin, N. Papernot, I. Goodfellow, D. Boneh, and P. McDaniel, "Ensemble adversarial training: Attacks and defenses," 2017, *arXiv:1705.07204*. [Online]. Available: https://arxiv.org/abs/1705.07204

[116] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 2107–2116.

[117] Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.

[118] H. Liu and T. Taniguchi, "Feature extraction and pattern recognition for human motion by a deep sparse autoencoder," in *Proc. IEEE Int. Conf. Comput. Inf. Technol.*, Sep. 2014, pp. 173–181.

[119] Y. Zeinali and B. A. Story, "Competitive probabilistic neural network," *Integr. Comput.-Aided Eng.*, vol. 24, no. 2, pp. 105–118, 2017.

[120] X. Miao, H. Miao, Y. Jia, and Y. Guo, "Using a stacked-autoencoder neural network model to estimate sea state bias for a radar altimeter," *PLoS ONE*, vol. 13, no. 12, 2018, Art. no. e0208989.

[121] W. Li, H. Fu, L. Yu, P. Gong, D. Feng, C. Li, and N. Clinton, "Stacked Autoencoder-based deep learning for remote-sensing image classification: A case study of African land-cover mapping," *Int. J. Remote Sens.*, vol. 37, no. 23, pp. 5632–5646, 2016.

[122] J. C. Junqua and J. P. Haton, *Robustness in Automatic Speech Recognition: Fundamentals and Applications*, vol. 341. New York, NY, USA: Springer, 2012.

[123] R. Sarikaya, G. E. Hinton, and A. Deoras, "Application of deep belief networks for natural language understanding," *IEEE/ACM Trans. Audio, Speech Language Process.*, vol. 22, no. 4, pp. 778–784, Apr. 2014.

[124] A. E. Omolara, A. Jantan, O. L. Abiodun, K. V. Dada, H. Arshad, and E. Emmanuel, "A deception model robust to eavesdropping over communication for social network systems," *IEEE Access*, vol. 7, pp. 100881–100898, 2019.

[125] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proc. 25th Int. Conf. Mach. Learn.*, Jul. 2008, pp. 160–167.

[126] H. Li, J. G. Ellis, L. Zhang, and S. F. Chang, "PatternNet: Visual pattern mining with deep neural network," in *Proc. ACM Int. Conf. Multimedia Retr.*, Jun. 2018, pp. 291–299.

[127] A. Bhandare, M. Bhide, P. Gokhale, and R. Chandavarkar, "Applications of convolutional neural networks," *Int. J. Comput. Sci. Inf. Technol.*, vol. 7, no. 5, pp. 2206–2215, 2016.

[128] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks," in *Proc. 23rd Int. Conf. Mach. Learn.*, Jun. 2006, pp. 369–376.

[129] Q. Wu, K. Ding, and B. Huang, "Approach for fault prognosis using recurrent neural network," *J. Intell. Manuf.*, to be published.

[130] M. Z. Uddin, W. Khaksar, and J. Torresen, "Facial expression recognition using salient features and convolutional neural network," *IEEE Access*, vol. 5, pp. 26146–26161, 2017.

[131] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, U.K.: MIT Press, 2016.

[132] M. Mamuda and S. Sathasivam, "Predicting the survival of diabetes using neural network," in *Proc. AIP Conf.*, vol. 1870, no. 1, 2017, Art. no. 040046.

[133] K. Y. Lee, N. Chung, and S. Hwang, "Application of an artificial neural network (ANN) model for predicting mosquito abundances in urban areas," *Ecol. Inform.*, vol. 36, pp. 172–180, Nov. 2016.

[134] D. Nofriansyah, Z. Effendi, and M. Ramadhan, "Application of discrete hopfield algortihm type artificial neural network for pattern recognition of eye's iris," *J. Saintikom*, vol, 16, no. 1, 2017.

[135] F. Cheng and V. Sutariya, "Applications of artificial neural network modeling in drug discovery," *Clin. Exp. Pharmacol.*, vol. 2, no. 3, pp. 1–2, 2012.

[136] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *J. Pharmaceutical Biomed.*, vol. 22, no. 5, pp. 717–727, 2000.

[137] C. I. Christodoulou and C. S. Pattichis, "Unsupervised pattern recognition for the classification of EMG signals," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 169–178, Feb. 1999.

[138] A. R. Syafeeza, M. Khalil-Hani, H. Imran, M. M. Ibrahim, and Y. C. Wong, "Generalizing convolutional neural networks for pattern recognition tasks," *ARPN J. Eng. Appl. Sci.*, vol. 10, no. 12, pp. 5298–5308, 2006.

[139] D. Siganos and C. Stergiou, "Neural networks, the human brain, and learning," Surveys Presentations Inf. Syst. Eng., Imperial College London, London, U.K., Tech. Rep., 1996.

[140] A. Dosovitskiy, J. T. Springenberg, M. Tatarchenko, and T. Brox, "Learning to generate chairs, tables and cars with convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 692–705, Apr. 2017.

[141] A. Bunte, A. Prahlad, and B. Brockway, "Systems and methods for creating copies of data, such as archive copies," U.S. Patent 8 140 786, Mar. 20, 2012.

[142] M. H. Hassoun, *Fundamentals of Artificial Neural Networks*. New York, NY, USA: IEEE Press, 1996,

[143] A. Zell, G. Mamier, M. Vogt, K. U. Herrmann, T. Soyez, M. Schmalzl, and B. Kett, *SNNS. Stuttgart Neural Network Simulator*, vol. 4. Cambridge, MA, USA: MIT Press, 1995.

[144] X.-H. Song and P. K. Hopke, "Kohonen neural network as a pattern recognition method based on the weight interpretation," *Anal. Chim. Acta*, vol. 334, nos. 1–2, pp. 57–66, 1996.

[145] N. Metawa, M. Elhoseny, A. E. Hassanien, and M. K. Hassan, *Expert Systems in Finance*. New York, NY, USA: Taylor & Francis, 2019.

[146] R. Elhassan and M. Ahmed, "Arabic text classification on full word," *Int. J. Comput. Sci. Softw. Eng.*, vol. 4, no. 5, pp. 114–120, 2015.

[147] M. H. Ghaseminezhad and A. Karami, "A novel self-organizing map (SOM) neural network for discrete groups of data clustering," *Appl. Soft Comput.*, vol. 11, no. 4, pp. 3771–3778, 2011.

[148] O. Vinyals, I. Babuschkin, J. Chung, M. Mathieu, M. Jaderberg, W. M. Czarnecki, and T. Ewalds, "AlphaStar: Mastering the real-time strategy game StarCraft II," *DeepMind Blog.*, 2019.

[149] Z. Dai, Z. Yang, Y. Yang, W. W. Cohen, J. Carbonell, Q. V. Le, and R. Salakhutdinov, "Transformer-xl: Attentive language models beyond a fixed-length context," 2019, *arXiv:1901.02860*. [Online]. Available: https://arxiv.org/abs/1901.02860

[150] K. Arulkumaran, A. Cully, and J. Togelius, "AlphaStar: An evolutionary computation perspective," 2019, *arXiv:1902.01724*. [Online]. Available: https://arxiv.org/abs/arXiv:1902.01724

[151] D. Heaven, "Two minds are better than one," *New Scientist*, vol. 243, no. 3244, pp. 38–41, 2019.

[152] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*. [Online]. Available: https://arxiv.org/abs/1810.04805

[153] Y. Deng, F. Bao, Y. Kong, Z. Ren, and Q. Dai, "Deep direct reinforcement learning for financial signal representation and trading," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 653–664, 2017.

[154] M. Göçken, M. Özçalici, A. Boru, and A. T. Dosdogru, "Integrating metaheuristics and artificial neural networks for improved stock price prediction," *Expert Syst. Appl.*, vol. 44, pp. 320–331, Feb. 2016.

[155] M. Qiu, Y. Song, and F. Akagi, "Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market," *Chaos, Solitons Fractals*, vol. 85, pp. 1–7, Apr. 2016.

[156] C. N. W. Tan, "A hybrid financial trading system incorporating chaos theory, statistical and artificial intelligence/soft computing methods," in *Proc. Queensland Finance Conf.*, 1999.

[157] W. Dai and P. Wang, "Application of pattern recognition and artificial neural network to load forecasting in electric power system," in *Proc. 3rd Int. Conf. Natural Comput. (ICNC)*, vol. 1, Aug. 2007, pp. 381–385.

[158] M. Inubushi and K. Yoshimura, "Reservoir computing beyond memory-nonlinearity trade-off," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 10199.

[159] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, "Recent advances in physical reservoir computing: A review," *Neural Netw.*, to be published.

[160] K. Abhishek, A. Kumar, R. Ranjan, and S. Kumar, "A rainfall prediction model using artificial neural network," in *Proc. IEEE Control Syst. Graduate Res. Colloquium (ICSGRC)*, Jul. 2016, pp. 82–87.

[161] R. R. Swain, P. M. Khilar, and T. Dash, "Multifault diagnosis in WSN using a hybrid metaheuristic trained neural network," *Digit. Commun. Netw.*, to be published.

[162] X. Y. Chen and K. W. Chau, "A hybrid double feedforward neural network for suspended sediment load estimation," *Water Resour. Manage.*, vol. 30, no. 7, pp. 2179–2194, 2016.

[163] M. A. Ghorbani, R. Kazempour, K. W. Chau, S. Shamshirband, and P. T. Ghazvinei, "Forecasting pan evaporation with an integrated artificial neural network quantum-behaved particle swarm optimization model: A case study in Talesh, Northern Iran," *Eng. Appl. Comput. Fluid Mech.*, vol. 12, no. 1, pp. 724–737, 2018.

[164] S. N. Qasem, S. Samadianfard, S. Kheshtgar, S. Jarhan, O. Kisi, S. Shamshirband, and K. W. Chau, "Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates," *Eng. Appl. Comput. Fluid Mech.*, vol. 13, no. 1, pp. 177–187, 2019.

[165] M. H. Ahmadi, B. Mohseni-Gharyehsafa, M. Farzaneh-Gord, R. D. Jilte, R. Kumar, and K. W. Chau, "Applicability of connectionist methods to predict dynamic viscosity of silver/water nanofluid by using ANN-MLP, MARS and MPR algorithms," *Eng. Appl. Comput. Fluid Mech.*, vol. 13, no. 1, pp. 220–228, 2019.

[166] C. L. Wu and K. W. Chau, "Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis," *J. Hydrol.*, vol. 399, nos. 3–4, pp. 394–409, 2011.

[167] P. G. Taherei, H. H. Darvishi, A. Mosavi, K. B. W. Yusof, M. Alizamir, S. Shamshirband, and K. W. Chau, "Sugarcane growth prediction based on meteorological parameters using extreme learning machine and artificial neural network," *Eng. Appl. Comput. Fluid Mech.*, vol. 12, no. 1, pp. 738–749, 2018.

[168] R. B. Abidoye and A. P. C. Chan, "Valuers' receptiveness to the application of artificial intelligence in property valuation," *Pacific Rim Property Res. J.*, vol. 23, no. 2, pp. 175–193, 2017.

[169] S. Vijayanarasimhan and K. Grauman, "Large-scale live active learning: Training object detectors with crawled data and crowds," *Int. J. Comput. Vis.*, vol. 108, nos. 1–2, pp. 97–114, 2014.

[170] Y. Dong, H. Su, J. Zhu, and F. Bao, "Towards interpretable deep neural networks by leveraging adversarial examples," 2017, *arXiv:1708.05493*. [Online]. Available: https://arxiv.org/abs/1708.05493

[171] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu, "Towards better analysis of deep convolutional neural networks," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 91–100, Aug. 2017.

[172] K. Kelchtermans and T. Tuytelaars, "How hard is it to cross the room?—Training (recurrent) neural networks to steer a UAV," 2017, *arXiv:1702.07600*. [Online]. Available: https://arxiv.org/abs/1702.07600

[173] R. Adams, Z. Ghahramani, and M. Jordan, "Tree-structured stick breaking processes for hierarchical data," 2009, *arXiv:1006.1062*. [Online]. Available: https://arxiv.org/abs/1006.1062

[174] M. F. Balcan and H. Zhang, "S-concave distributions: Towards broader distributions for noise-tolerant and sample-efficient learning algorithms," *Stat*, vol. 1050, p. 22, Mar. 2017.

[175] V. C. Cardei, B. Funt, and K. Barnard, "Adaptive illuminant estimation using neural networks," In *Proc. ICANN*. London, U.K.: Springer, 1998, pp. 749–754.

[176] P. A. Ryan, Z. Ghahramani, and M. I. Jordan, "Tree-structured stick breaking for hierarchical data," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 23, 2010, pp. 19–27.

[177] J. H. Murphy and B. J. Kagle, "Neural network recognition of electronic malfunctions," *J. Intell. Manuf.*, vol. 3, no. 4, pp. 205–216, 1992.

[178] M. Vellanki and C. H. Dagli, "Artificial neural network approach in printed circuit board assembly," *J. Intell. Manuf.*, vol. 4, no. 1, pp. 109–119, 1993.

[179] H. Jack, D. M. A. Lee, R. O. Buchal, and W. H. Elmaraghy, "Neural networks and the inverse kinematics problem," *J. Intell. Manuf.*, vol. 4, no. 1, pp. 43–66, 1993.

[180] D. M. D'Addona, A. S. Ullah, and D. Matarazzo, "Tool-wear prediction and pattern-recognition using artificial neural network and DNA-based computing," *J. Intell. Manuf.*, vol. 28, no. 6, pp. 1285–1301, 2017.

[181] S. B. Brahim, A. E. Smith, and B. Bidanda, "Relating product specifications and performance data with a neural network model for design improvement," *J. Intell. Manuf.*, vol. 4, no. 6, pp. 367–374, 1993.

[182] T. Van Tung and B.-S. Yang, "Machine fault diagnosis and prognosis: The state of the art," *Int. J. Fluid Mach. Syst.*, vol. 2, no. 1, pp. 61–71, 2009.

[183] J. Liu, X. Kong, F. Xia, X. Bai, L. Wang, Q. Qing, and I. Lee, "Artificial intelligence in the 21st century," *IEEE Access*, vol. 6, pp. 34403–34421, 2018.

[184] M. C. Arranz, W. Birk, and G. Nikolakopoulos, "A survey on control configuration selection and new challenges in relation to wireless sensor and actuator networks," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8810–8825, 2017.

[185] C. Kanellakis and G. Nikolakopoulos, "Survey on computer vision for UAVs: Current developments and trends," *J. Intell. Robot. Syst.*, vol. 87, no. 1, pp. 141–168, Jul. 2017.

[186] D. Lee, N. He, and J. Hu, "Dynamic programming for stochastic control systems with jointly discrete and continuous state-spaces," 2018, *arXiv:1803.08876*. [Online]. Available: https://arxiv.org/abs/1803.08876

[187] L. H. Thai, N. D. T. Nguyen, and T. S. Hai, "A facial expression classification system integrating canny, principal component analysis and artificial neural network," 2011, *arXiv:1111.4052*. [Online]. Available: https://arxiv.org/abs/1111.4052

[188] J. Mahanta, "Introduction to neural networks, advantages and applications," *Towards Data Sci.*, 2017.

[189] M. Edoardo, T. T. Long, and R. J. Nicholas, "On the efficiency of data collection for crowdsourced classification," in *Proc. IJCAI*, 2018, pp. 1568–1575.

[190] J. Wilder, S. Dickinson, A. Jepson, and D. B. Walther, "Spatial relationships between contours impact rapid scene classification," *J. Vis.*, vol. 18, no. 8, p. 1, 2018.

[191] J. Fürnkranz, "Separate-and-conquer rule learning," *Artif. Intell. Rev.*, vol. 13, no. 1, pp. 3–54, 1999.

[192] J. K. Tarus, Z. Niu, and G. Mustafa, "Knowledge-based recommendation: A review of ontology-based recommender systems for e-learning," *Artif. Intell. Rev.*, vol. 50, no. 1, pp. 21–48, Jun. 2018.

[193] R. J. Nicholas, "On the application and future of agents and multi-agent systems," *KI*, vol. 18, no. 2, p. 38, 2004.

[194] R. J. Nicholas, R. Alex, S. Case, R. Johnston, and D. Philpot, "Decentralized data and information systems: Theory and practice," *Comput. J.*, vol. 53, no. 9, pp. 1341–1343, 2010.

[195] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, "Applications of deep learning and reinforcement learning to biological data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2063–2079, Jun. 2018.

[196] B. M. Shah, "Improving anomaly detection process in computer networks having existing IDs using additional behavioral layer optimized back propagation neural network and mobile agents multi class attack detection," Ph.D. dissertation, Dept. Comput. Sci., Gujarat Technol. Univ., Ahmedabad, India, 2016.

[197] S. Chen, J. Yang, L. Luo, Y. Wei, K. Zhang, and Y. Tai, "Low-rank latent pattern approximation with applications to robust image classification," *IEEE Trans. Image Process.*, vol. 26, no. 11, pp. 5519–5530, Nov. 2017.

[198] A. Jain and A. M. Kumar, "Hybrid neural network models for hydrologic time series forecasting," *Appl. Soft Comput.*, vol. 7, no. 2, pp. 585–592, Mar. 2007.

[199] G. T. Tsenov and V. M. Mladenov, "Speech recognition using neural networks," in *Proc. 10th Symp. Neural Netw. Appl. Elect. Eng. (NEUREL)*, 2010, pp. 181–186.

[200] T. E. Saaed, G. Nikolakopoulos, J. E. Jonasson, and H. Hedlund, "A state-of-the-art review of structural control systems," *J. Vib. Control*, vol. 21, no. 5, pp. 919–937, 2015.

[201] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemometrics Intell. Lab. Syst.*, vol. 39, no. 1, pp. 43–62, 1997.

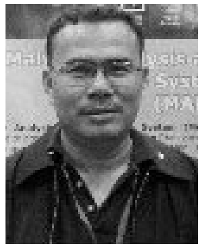
[202] Y. Upadhyay, "Introduction to feedforward neural networks," *Towards Data Science*, vol. 7, 2019.

[203] M. T. Khadir and J. Ringwood, "Neural network modelling and predictive control of a milk pasteurisation plant," Dept. Electron. Eng., National Univ. Ireland, Maynooth, Ireland, 2001.

[204] J. Lian, W. Zhang, J. Hu, and S. H. Zak, "Variable neural adaptive robust output feedback control of uncertain systems," in *Proc. Amer. Control Conf. (ACC)*, Jun. 2010, pp. 99–104.

[205] A. Esparza, A. Sala, and P. Albertos, "Neural networks in virtual reference tuning," *Eng. Appl. Artif. Intell.*, vol. 24, no. 6, pp. 983–995, 2011.
[206] A. A. Eielsen, J. T. Gravdahl, and K. Y. Pettersen, "Adaptive feed-forward hysteresis compensation for piezoelectric actuators," *Rev. Sci. Instrum.*, vol. 83, no. 8, 2012, Art. no. 085001.
[207] F. Khorrami, S. Jain, and A. Tzes, "Experimental results on adaptive nonlinear control and input preshaping for multi-link flexible manipulators," *Automatica*, vol. 31, no. 1, pp. 83–97, 1995.
[208] C. Gelmereanu, L. Morar, and S. Bogdan, "Productivity and cycle time prediction using artificial neural network," *Procedia Econ. Finance*, vol. 15, pp. 1563–1569, Jan. 2014.
[209] S. Y. Yun, S. Namkoong, J. H. Rho, S. W. Shin, and J. U. Choi, "A performance evaluation of neural network models in traffic volume forecasting," *Math. Comput. Model.*, vol. 27, nos. 9–11, pp. 293–310, 1998.
[210] H. M. Elragal, "Improving neural networks prediction accuracy using particle swarm optimization combiner," *Int. J. Neural Syst.*, vol. 19, no. 5, pp. 387–393, 2009.
[211] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, vol. 2, Jul. 2004, pp. 985–990.
[212] I. Marović, I. Sušanj, and N. Ožanić, "Development of ANN model for wind speed prediction as a support for early warning system," *Complexity*, vol. 2017, Dec. 2017, Art. no. 3418145.
[213] R. Peterson, A. Baldwin, and A. Schmidt, "Comparative analysis of backpropagation algorithm against the ant colony optimization," *Int. J. Comput. Sci.*, vol. 1, no. 3, pp. 83–88, 2019.
[214] S.-H. Oh, "Improving the error backpropagation algorithm with a modified error function," *IEEE Trans. Neural Netw.*, vol. 8, no. 3, pp. 799–803, May 1997.
[215] M. Puig-Arnavat and J. C. Bruno, "Artificial neural networks for thermochemical conversion of biomass," in *Recent Advances in Thermo-Chemical Conversion of Biomass*. Amsterdam, The Netherlands: Elsevier, 2015, pp. 133–156.
[216] G. N. Ezeh, N. S. Chukwuneke, N. C. Ogujiofor, and U. H. Diala, "Effects of rain attenuation on satellite communication link," *Adv. Sci. Technol. Res. J.*, vol. 8, no. 22, pp. 1–11, 2014.
[217] A. El-Shahat, *Advanced Applications for Artificial Neural Networks*. Norderstedt, Germany: Books on Demand, 2018.
[218] M. P. Brenner, L. S. Levitov, and E. O. Budrene, "Physical mechanisms for chemotactic pattern formation by bacteria," *Biophys. J.*, vol. 74, no. 4, pp. 1677–1693, 1998.
[219] R. Cooper, S. Edgett, and E. Kleinschmidt, "Portfolio management for new product development: Results of an industry practices study," *R&D Manage.*, vol. 31, no. 4, pp. 361–380, 2001.
[220] S. Pantawane, D. Bandyopadhyay, and N. Choudhury, "Generic mechanism for pattern formation in the solvation shells of buckminsterfullerene," *ACS Omega*, vol. 3, no. 1, pp. 1060–1068, 2018.
[221] M. Oschatz and R. Walczak, "Crucial factors for the application of functional nanoporous carbon-based materials in energy and environmental applications," *C*, vol. 4, no. 4, p. 56, 2018.
[222] M. Dodgson, *Technological Collaboration in Industry: Strategy, Policy and Internationalization in Innovation*. Evanston, IL, USA: Routledge, 2018.
[223] T. Zan, Z. Liu, H. Wang, M. Wang, and X. Gao, "Control chart pattern recognition using the convolutional neural network," *J. Intell. Manuf.*, to be published.
[224] A. Addeh, A. Khormali, and N. A. Golilarz, "Control chart pattern recognition using RBF neural network with new training algorithm and practical features," *ISA Trans.*, vol. 79, pp. 202–216, Aug. 2018.
[225] Y. E. Shao, P.-Y. Chang, and C.-J. Lu, "Applying two-stage neural network based classifiers to the identification of mixture control chart patterns for an SPC-EPC process," *Complexity*, vol. 2017, Oct. 2017, Art. no. 2323082.
[226] S. Sağiroğlu, E. Be dok, and M. Erler, "Control chart pattern recognition using artificial neural networks," *Turkish J. Elect. Eng. Comput. Sci.*, vol. 8, no. 2, pp. 137–146, 2017.
[227] L. S. Nelson, "The Shewhart control chart—Tests for special causes," *J. Qual. Technol.*, vol. 16, no. 4, pp. 237–239, 1984.
[228] I. Masood and A. Hassan, "Issues in development of artificial neural network-based control chart pattern recognition schemes," *Eur. J. Sci. Res.*, vol. 39, no. 3, pp. 336–355, 2010.

OLUDARE ISAAC ABIODUN received the Ph.D. degree in nuclear and radiation physics from the Nigerian Defence Academy, Kaduna, and the Ph.D. degree in computer science and security from Universiti Sains Malaysia, Penang, Malaysia. His research interests include artificial intelligence, robotics, cybersecurity, digital forensics, nuclear security, terrorism, and national security.

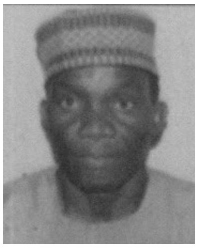


AMAN JANTAN is currently an Associate Professor with the School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia. He has published more than one hundred articles in reputable journals. His research interests include digital forensic, artificial intelligence, Malware, intrusion detection systems, computer security, cryptography, and computer and network security. He received national and international recognition in some of his work.

ABIODUN ESTHER OMOLARA is currently pursuing the Ph.D. degree with the School of Computer Sciences, Universiti Sains Malaysia. Her research interests include computer and network security, digital forensic, cyber-security, cryptography, information assurance, artificial intelligence, and natural language processing.

KEMI VICTORIA DADA received the B.Sc. degree from Nasarawa State University, Keffi, Nigeria, and the M.Sc. degree from Ahmadu Bello University, Zaria, Nigeria, all in statistic. Her research interests include statistics, economic and forensic analysis, security information analysis, cybercrime solution, and computational intelligent.

ABUBAKAR MALAH UMAR received the Ph.D. degree in nuclear and radiation physics from the Nigerian Defence Academy, Kaduna. He was with the Energy Commission of Nigeria. He is currently the Director and the Head of the Department Nuclear Science and Technology. His research interests include nuclear power plant, nuclear security, nuclear radiation safety, solar power, artificial intelligence, and computer science.



OKAFOR UCHENWA LINUS received the Ph.D. degree in applied mathematics from the Nigerian Defence Academy, Kaduna. He is currently a Senior Lecturer. His research interests include mathematics, statistics, modelling, simulations, numerical analysis, optimization, physics, artificial intelligence, computer science, artificial intelligence, robotics, data mining, information security, intrusion detection, reverse engineering, and the semantic web.

HUMAIRA ARSHAD received the Ph.D. degree in computer sciences from Universiti Sains Malaysia, Penang, Malaysia. She is currently a Lecturer with the Department of Computer Science and Information Technology, Islamia University of Bahawalpur, Pakistan. Her research interests include neural networks, cybersecurity and digital forensics, information security, online social networks, social engineering, data mining computer, and artificial intelligence.

ABDULLAHI AMINU KAZAURE is currently pursuing the Ph.D. degree in digital forensics with the School of Computer Sciences, Universiti Sains Malaysia. He is currently a Lecturer with Hussaini Adamu Federal Polytechnic Kazaure, Nigeria. He is also a Full Member of the Nigerian Computer Society (NCS). His research interests include artificial intelligence, digital forensics, cybersecurity, networking, and intrusion detection.

USMAN GANA received the B.Sc. degree in computer science from Usman Dan Fodiyo University, Sokoto, in 2006, and the M.Sc. in computer science from the University of Greenwich London, U.K., in 2010. He is currently pursuing the Ph.D. degree with Universiti Sains Malaysia. He has been a Lecturer I with Ibrahim Badamasi Babangida University, Lapai, Nigeria, since 2011. His research interests include cloud computing, network security, cyber security, and green computing.

MUHAMMAD UBALE KIRU received the bachelor's degree in English and the master's degree in information technology from Lovely Professional University, Phagwara, in India. He is currently pursuing the Ph.D. degree with the School of Computer Science, Universiti Sains Malaysia, Penang, Malaysia. He has other certifications, such as Microsoft's MCP, MCSA, MCITP, Cisco's CCNA, EC-Council's CEH and CompTIA Security+. His research interests include machine learning, malware forensics, server administration, ethical hacking, vulnerability assessment and cloud storage security, and network management. He has authored a few research articles in the last few years.