Assistive Technology Evolving as Intelligent System

Amlan Basu [1], Lykourgos Petropoulakis [2], Gaetano Di Caterina [3], and John Soraghan [4]

[1] Ph.D. Student, CeSIP. Royal College Building, University of Strathclyde, Glasgow, U.K.
amlan.basu@strath.ac.uk

[2] Senior Knowledge Exchange Fellow, CeSIP. Royal College Building, University of Strathclyde, Glasgow, U.K.
l.petropoulakis@strath.ac.uk

[3] Research Fellow, CeSIP. Royal College Building, University of Strathclyde, Glasgow, U.K.
gaetano.di-caterina@strath.ac.uk

[4] Professor, CeSIP. Royal College Building, University of Strathclyde, Glasgow, U.K.
j.sorghan@strath.ac.uk

Abstract - Different evolving technologies surround humans today. Among the various technologies, Assistive Technology has still not established itself firmly because there is an absence of proper integration of this technology with human life. However, in the future, it will become one of the most important and vital phenomena in everyone's life. Because humans want to make their life easier and longer and these are the reasons for the rapid growth in demand for Assistive Technology. Therefore, improvements in the technology and the way it is applied are essential and, for this reason, there is a requirement of a detailed study of the technology. This paper demonstrates the different milestones achieved in assistive technology by using different techniques to attempt to improve intelligence in assistive systems; and also, it describes the gaps that are still present even after such extensive works and, which are required to be either resolved or bridged. This study is done to understand where the assistive technology is today and in which direction it needs to get directed.

Keywords – Assistive Technology, Artificial Neural Networks, Artificial Intelligence, Machine Learning, Intelligent systems

1. Introduction

Assistive Technology has started evolving drastically. Usage of the term assistive system is for something that assists humans or living beings in some way or another. For example, when a person needs to follow a daily routine as prescribed by the doctor to remain healthy, then this person needs to track everything. He or she has to remember which medicine to take and when to go for certain health-related tests and also when there is an appointment with the doctor. Therefore, to make this task simple an assistive system in the form of a talking robot has been developed and it has been named as Pillo. Pillo keeps track of everything, like providing the person with medicine on time, tracking the progress of calories burned by the person during exercise and helping the person to talk to the doctor through video conferencing whenever there is a need or an appointment.

There is another robot assistive system invented which is known as Aido. It is a multi-tasking robot. In the morning it wakes up every person at home, at different times prescribed by the persons relying on Aido. It can navigate inside the house and is also capable of monitoring the security of the house by recognizing the faces of the people. Also, if a person is not at home, he/she can see and talk to any person visiting the home with the help of Aido.

The robot helps children in education and related knowledge assistance. There are many other intelligent assistive systems like Moorebot, Miko, etc. which are also used inside the home for different purposes.

So, the whole concept of developing an assistive system has evolved from simple computer systems to robots or personal assistants. Now the assistive technology is evolving to become an intelligent one (an intelligent system is a system that performs all the assigned tasks on its own) because of the endless demands by humans, who always remain unsatisfied as they always want more improvements and better functionality. Therefore, the concepts of artificial intelligence (neural network, machine learning, etc.) are used to develop the assistive system.

Assistive Technology has made its presence felt in every sector of the world, right from the health-related issues to education. Detection and recognition of objects, face and facial expression recognition and detection, assisting people in remaining physically fit are some of the systems developed using various artificial intelligence techniques until today [1].

Before moving into the discussion of different works done in this sector. There will be an explanation of the various techniques used to make the systems intelligent. Then there will be a discussion on the different essential and consequential works, and at last, there will be a discussion on various pros and cons related to the sector and also on the solutions associated with the problems which can help to bridge the gap or resolve them.

2. Artificial Intelligence

Artificial Intelligence (AI) is a method of developing a system that can act rationally and also like a human. Moreover, a system that can even think logically and can perform the tasks like humans without taking any support form humans. Furthermore, it can be something that is more precise as compared to human because of the storage capacity and learning ability [2].

Artificial Intelligence involves fuzzy logic, genetic algorithm, neural networks, etc. The use of machine learning or deep learning concepts is for training the developed system. So, the concepts present in AI are responsible for the development of intelligent systems or making an existing system intelligent.

In the coming sections, some of the different areas associated with AI are discussed. After the discussion on areas related to AI, the discussion gradually enters the intelligent systems developed using some concepts of AI.

3. Fuzzy Logic

In traditional logic, there are only two answers to any given problem that is true and false. However, this logic is not applicable in fuzzy logic, the answer in the fuzzy logic can lie between the true and the false or in other words the answer to the problem is partially true or partially false. For instance, the example of hot and cold water, in traditional logic the answer to the question of whether the water is hot or cold will be either hot or cold whereas in fuzzy logic it will be a degree to which the water is cold or hot.

Fuzzy logic is very useful in the situations where the answers with only yes or no can be confusing and dangerous. For example, in automatic braking systems of a car. Using the brakes only when the car is close to some object and do not use the brakes when it is far. However, this logic is absurd and dangerous. When the same problem is solved using fuzzy logic, then it gives a tentative value that lies between 0 (false/no) and 1 (true/yes). Thus, there is a use of brakes with varied pressure on it [3].

Unlike traditional controllers, fuzzy controllers do not require a complete specification of a system, and fuzzy controllers are robust as compared to traditional controllers. The requirement

is only of a mathematical expression that helps the controller to deal with the different inputs and produce a suitable output.

The application of fuzzy logic is mainly related to control systems like kitchen, car, aerospace vehicles, etc. In addition to this, robotics, decision making, route-planning, image processing, etc. are the areas where fuzzy logic plays an important role. However, it would be wrong to say that applying fuzzy logic in a system makes it intelligent. Because an intelligent system can learn from the provided dataset or by seeing the things around it, and this phenomenon is not happening in the case of fuzzy logic. In fuzzy logic, a particular developed mathematical methodology is fed into the system to obtain the desired outputs. For example, an air conditioner (AC). AC is semi-automatic when the users adjust its temperature according to how the user feels at the different time. The same is automatic when the system decides on when to cool the area and when to stop when the area is cold, and the same system is fully-automatic when it can adjust its temperature by taking the reference of temperature in its surrounding. This system is sometimes also termed as intelligent, but this system is partially intelligent as it can't learn from different experiences, rather it is the already fed in methodology or formula that is helping to make some decision by external inputs received.

Using the example of robotic kitchen hands that can cook, also can help us understand this phenomenon. Earlier the robotic hand in the kitchen could cook only those foods whose recipes are already present in it, and used fuzzy logic to operate different things like controlling the heating, grabbing the utensils, etc. [4].

Now the same hand doesn’t need any recipes stored in it, rather it learns from the cook by just observing on how the cook prepared the food and then emulates everything and prepares the same food. This kind of system is known to be intelligent. Of course in this system as well there is the application of fuzzy logic but with an additional feature of neural network architecture that makes the system capable of learning.

The fuzzy logic methodology is not the way in which a system is fed with a huge chunk of the dataset to learn about different things and produce a result by the learning experience. In fact, fuzzy logic is a methodology applied in a system that can process the input to produce the user’s desired output. Thus, fuzzy logic can help us get a system that is fully-automatic or a system that is partially intelligent. However, it is not capable of producing a complete intelligent system on its own. Therefore, the contribution of fuzzy logic in creating an intelligent system is invaluable, but itself can’t create an intelligent system, rather it can help in producing outputs that can help the system to learn and take experience and with the latter can produce results.

The point to remember is that the concept of fuzzy logic and fractional systems are very different. Fuzzy logic is dependent on set theory whereas fractional systems completely rely on concepts of fractional calculus [5]. Fuzzy logic a way to embed a methodology in one of the systems to get the desired output and fuzzy logic help in developing controllers as well. However, fractional systems [6] are helpful to understand the exact behavior of a system. A system can be of integer form or fractional form. Integer order systems provide lesser information about a system as compared to the fractional order of the same system. Fractional logic can help in developing suitable controllers. However, they do not match the performance criteria and robustness provided by the controllers designed using fuzzy logic, neural networks, genetic algorithm, particle swarm optimization, etc.

4. Genetic Algorithm

Genetic Algorithm (GA) is a technique which is used to solve the problem of non-linearity using optimization. It uses an optimization process (or an iterative process) to reach the global minimum value by mimicking the biological concepts. The parent fields that have inspired GA to come into existence are the evolutionary algorithms, metaheuristics, stochastic optimization,

optimization and evolutionary computing. GA uses the four basic steps in optimization, initialization, selection, crossover, and mutation. GA rapidly identifies the changes that are happening and is very much useful in searching for larger problems.

For the optimization GA requires two basic things, the first is a genetic representation of any problem, i.e., the representation of the chromosomes in a proper way and the second is a fitness function which is a mathematical expression (an objective function) that helps to find out the efficiency of a solution to achieve the set aims [7].

GA has played a vital role in revolutionizing the image processing areas, control systems, quality control, etc. However, the use of GA in neural networks for making the system intelligent is very less. Usage of GA in a neural network is seen only in Recurrent Neural Network (RNN) and to obtain classified examples for training the neural networks.

There are some serious disadvantages to GA. Finding accurate global maxima is not guaranteed. The time taken for convergence is too long. It is a trial and error method, i.e.; it can't be said that which particular values for every step will produce the best result. Choosing the fitness function, genetic encoding, etc. are also important as they play a vital role in system effectiveness. And the last is that even after following so many things to get an optimized value it cannot be assured that one can get a comprehensible solution.

However, to overcome the disadvantages present in GA, different other ways of optimization and techniques are present. Some of the techniques that overcome the problems present in GA are swarm intelligence or SI (Particle Swarm Optimization or PSO is one of the well known SI and highly used technique, ant colony, and stochastic diffusion search), hill climbing, gaussian adaptation, etc.

Therefore, GA, SI, etc. alone cannot make a system cmpletely intelligent. Rather it can be said that they make a system partially intelligent. They play a vital role in the development of AI as a section to generate some desired outputs so that the generated outputs act as input for another section that is responsible for making a system completely intelligent. This statement is relevant because GA, SI, etc. alone are not capable of giving the power to a system to learn from a given data or the power to learn on its own [8].

5. Expert Systems

Expert Systems (ES) in context of AI, is a system that behaves like a human expert but is a computer system. It first came into existence in 1970 because of Prof. Edward Feigenbaum. The first ES developed was the Lisp machine. ES has two very important parts, the database, and the inference engine. ES works on a very simple architecture and has a very simple programming concept. For instance, the automatic spelling check in Microsoft word is nothing but an ES. In Microsoft word when a word is typed wrong or instead of using a capital letter, the small letter is used then it automatically corrects it and this because of the ES mechanism. When a user types a word, then the inference engine reads that word and processes it such that it can compare it with a list of similar words present in the database (prepared by using the inputs from human experts). So, if it finds that the word is confusing but is almost similar to the word present in the database, then it automatically corrects it [9].

The use of ES mechanism is huge in the medical sector. Some of the countries have very less number of doctors. Such countries installed ES systems in which a patient can feed the problems, he/she is facing, and on that basis, the ES using its inference engine and database prepared to contain the inputs from medical experts prescribes the medicines and also medical tests required, to the patient. The advantage of the ES is that it provides service anytime the user wants, and it never expires with such a huge amount of database.

However, observations on ES suggest that it lacks common sense [10]. For example, the ES system for treating patients called MYCIN [11, 12] developed for treating bacterial diseases.

In MYCIN, the patient gives an input that he/she has *Cholerae vibrio*, then the ES quickly suggests tetracycline medicine for two weeks. But it is unable to understand that the patient will be dead before the course is complete because of cholera. Therefore, the ES only looks towards the specific bacteria but not to the problems associated with that particular disease. Moreover, criticism of ES also exists because of the high prices that make it unaffordable.

ES is an intelligent system or not remained a topic of debate until ROSS came into existence. ROSS is a highly developed ES that has the qualities of data mining, deep learning, pattern recognition and natural language processing ability. Using these tools, it acts as an attorney suggesting solutions and became the first AI attorney. Also, ROSS is the first intelligent ES that has an inference engine and also a database fed with solutions from different experts [13]. It is said to be the first intelligent ES because it will not only produce solution or corrections but also because of the presence of deep learning it will learn from the experience it is going to come across and that will help increase its database without the help of any external expert feeding data into it. Therefore, ROSS can develop its database on its own and can also accept data externally fed and stored in it. Therefore, the introduction of deep learning in ES may certainly solve the problem of common sense that persisted in it.

6. Artificial Neural Network (ANN)

The complex networks formed by joining different units (neurons) and where each unit is dependent on another are called artificial neural network. The ANN concept originated from the ideas of natural neural systems of humans. An ANN imitates the function of a biological neural network, and it forms a system similar to the biological neural network. These neural networks are trained for some particular task several times to perform the assigned work precisely. The more the training, the better the system will be [14].

Figure 1 shows the schematic of the ANN. From the figure, it is clear that an ANN has an input layer, an output layer, and hidden layers. There may be many inputs and also many outputs and also there may be N number of hidden layers. All hidden layers are connected to each input, and also each unit of the hidden layer is connected to each unit of another hidden layer. The number of hidden layers can vary from one to many. The larger the number of hidden layers the higher is the complexity and greater is the accuracy. When the number of hidden layers is many, then the networks are specified as deep neural networks (DNN). So, the figure thoroughly explains how an ANN looks like and how are the units (neurons) connected to each other [15].

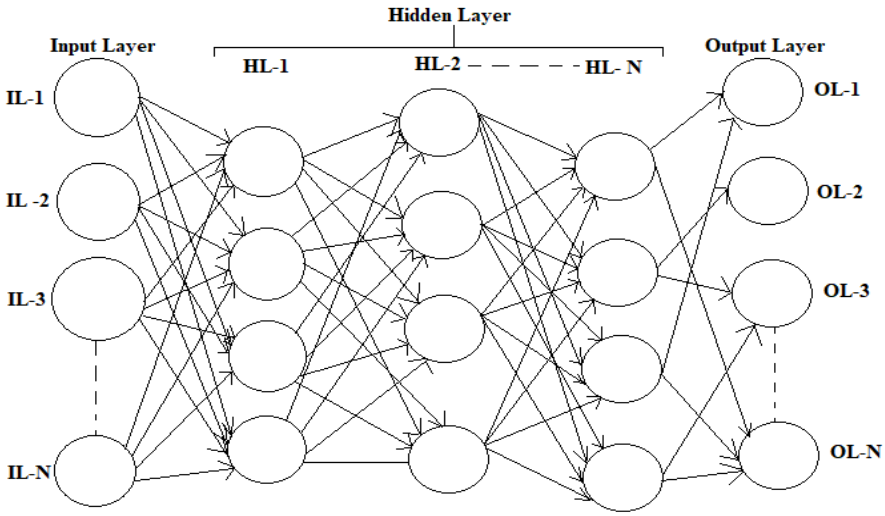

Fig. 1 A fully connected Artificial Neural Network (ANN) with N number of layers.

DNNs are the basic architecture because of which the other neural network in the world of deep learning (DL) [16] came into existence. When DNNs became incompetent then by making the DNNs the fundamental concept, other neural networks were proposed. The first concept proposed was the convolutional neural network (CNN) which is today the most basic architecture to form any advanced neural network for different purposes. The performance of CNNs is remarkable in various tasks, and this is because of the different architecture of CNN proposed over the years.

The first and the most basic architecture of CNN is the LeNet-5 by Yann LeCun et al. [17] in 1998. Then came the AlexNet in 2012 that was presented by Alex Krizhevsky et al. [18] and this proposed architecture of CNN got the first position in ILSVRC (ImageNet Large Scale Visual Classification) 2012. AlexNet was also introduced with batch normalization, and it was termed as Batch Normalized AlexNet (BN-AlexNet). In 2013, another CNN called ZFNet was proposed by Mathew Zeiler et al. [19] that acquired the first position in ILSVRC 2013. In 2014, VGG came into existence that was proposed by Karen Simonyan et al. [20]. There are two types of VGG present, VGG-16, and VGG-19. 16 and 19 are the number of layers present in the proposed CNN. Inspite of never have won in ILSRVC, it showed remarkable performance with a whole new concept of CNN architecture. In 2014, GoogLeNet proposed by Christian Szegedy et al. [21] that achieved the first position in ILSVRC 2014. Then in 2015 Kaiming he et al. [22, 23] proposed ResNet which won the ILSVRC 2015 and remains the best until now. ResNet also has a special module in its architecture, that makes it different from other CNN architectures, which is known as Inception module. There are two versions of inception module introduced till now, inception v3 [24] and inception v4 [25] and they both have helped increase the performance accuracy of ResNet.

As the work on CNN architecture progressed for attaining the highest achievable accuracy, the number of layers in it increased as well over the years. The LeNet-5 had only seven layers, AlexNet with eight layers, VGG with 16 and 19 layers, GoogleNet with 22 layers and ResNet with 18, 34, 50, 101 and 152 layers. The highest is the ResNet with 152 layers in 2015, and this coined a new terminology called Deep Convolutional Neural Networks (D-CNN). However, the basic problems related to detection and recognition still pursued. Therefore, other more sophisticated neural networks with CNN came into existence are discussed later in this paper.

7. Machine Learning

Machine learning (ML) is the way of developing a learning algorithm that can be embedded in a system to make it an intelligent one. ML has become very popular because of shorter programs that are easier to maintain and are more likely to be accurate. ML can also be used for solving many complex problems, which are difficult to solve using the traditional way. The same is evident in many of the cases like spam detection in e-mails, tagging photographs on Facebook, and so on. There are many types of ML which are classified into two categories, supervised learning and unsupervised learning.

In both supervised and unsupervised learning, there is a data set provided to the system. However, using the data set in supervised learning, the system is dictated or guided to produce the desired outputs, and in unsupervised learning, the system is left free to predict an output on its own.

Apart from these two, there are other types of semi-supervised learning, reinforcement learning, batch learning, online learning, instance-based learning, model-based learning, and recommender system. The definitions of each learning are as follows,

- Semi-supervised learning is a mixture of the supervised and unsupervised learning system.

- Reinforcement learning is a way in which a system learns by observing and analyzing the things happening and then doing the task with the help of experience it earned during observation.
- Batch learning, also known as offline learning, in which a system is provided data during its production as it cannot take any further data once its creation is complete, in other words, the system which cannot learn incrementally. If there is a need of implementation of new data, then a new version of the system is required to be developed.
- Online learning, also termed as incremental learning is just the opposite of what batch learning is. So, a system having the quality of online learning can take in new data in a particular sequence and can learn from them even after its production.
- Instance-based learning is the way in which the system learns from some examples and tries to predict the output for input by associating it with the examples that it already learned.
- Model-based learning is the way in which the system creates models of different examples, or in other words groups, the examples under different categories and then using these groups or models predicts a suitable output for the input [26].

Therefore, the concept of ML is the backbone for creating an intelligent system because it helps the system in developing a proper logic to acquire a particular way of learning different things and then act accordingly [27].

8. Important Developments in Intelligent Assistive Systems

Since all the basic terminologies and concepts that are necessary to be known to understand the development done in the field of Intelligent Assistive Systems (IAS), are well discussed. Therefore, this section contains a brief explanation and discussion on the various state-of-the-art works accomplished in the field of IAS.

The first one is the invention of a traffic signal recognition system that has an accuracy of 99.46%, which is more than that of human which is 98.84%. The achieved result is by forming the multicolumn deep neural network (MCDNN), which is nothing but an artificial neural network being developed using multiple deep neural networks for acquiring robustness. In this case, it was formed using 5 per pre-processing ways and 25 nets. Also, the error occurring in the recognition task reduced three times. The recognition rate increased from 98.52% to 99.46%. The trained MCDNN was tested practically and won the German traffic sign recognition benchmark [28, 29].

Another development had been for the visually impaired people. The system has been developed to provide information about people's surroundings. In this system a 360^0 camera is used, that captures the scene and sends it to a mobile device. To make the system intelligent the convolutional neural network (CNN) and an already trained VGG-19 (Visual Geometry Group -19) network had been used, to recognize the objects in the captured image. Then the complete message is given to the user (person) in an audio manner. The accuracy attained for this particular system is 92.8% [30].

A similar system was developed for visually impaired people such that they can easily navigate in indoor space. In this system, the home premises have installed IP cameras. The cameras capture the image and send it to the application installed on the smartphone used by the concerned person. The application using the image processing finds out the objects present in it and then informs the relevant person by producing an audio message [31].

Further, an intelligent program with very high precision, which was developed to play the game 'Go' (considered as a very complex game) and win it quite easily. 'AlphaGo' is the name

given to the developed program. The accuracy rate achieved in this was 99.8%, it defeated the professional player and champion of this game in all the games played. There are two necessary things in this development; the first one is the value network implemented for analyzing the position of the board and the second is policy networks, implemented for choosing the next move to make and these are added with Monte Carlo simulation which gave rise to a new search algorithm. The learning methods used for making the program intelligent are the supervised learning and reinforcement learning [32, 33].

The development of Faster-RCNN (Regional Convolutional Neural Network) done for the detection of real-time objects is an another finest development. This neural network is the amalgamation of RPN (Regional Proposal Network) and Fast-RCNN [34] (a neural network with the convolutional neural network, regions of interest, fully connected layers and followed with two outputs, softmax layer, and bounding-box regressor layer). After its testing, the speed reported was of 5 frames per second with an appreciable accuracy rate [35].

For the same purpose one more neural network has been proposed, Mask-RCNN. It is created by amalgamating the faster-RCNN and FCN (Fully Convolutional Networks) [36]. In this neural network architecture, the Faster-RCNN creates the bounding boxes whereas the RPN and FCN create the mask over the object and the object present in the image is detected. This architecture has proved to be one of the best in detecting the objects as it has a very high accuracy rate because of the introduction of RoI (region of interest) Align in the architecture, however, the speed is still the same as that of faster-RCNN. Moreover, this architecture had helped to solve the problem of instantaneous segmentation, as the architectures proposed before had only the ability of object detection and semantic segmentation [37].

YOLO (You Only Look Once) is a neural network that has been developed to track and detect different objects at a very fast pace. It has been specifically developed to meet the high-speed criteria. There are many versions of YOLO that are in existence, and their speed varies from 45 frames per second to 166 frames per second. Some of the versions of YOLO are, YOLO v1, Tiny YOLO, YOLO v2 and YOLO v3. According to the results obtained, YOLO works 1000 times faster than R-CNN and 100 times faster than Fast R-CNN [38-40].

Multi-Resolution Convolutional Neural Network (MRCNN) is responsible for large-scale image classification. For training the neural network, the MIT Places dataset [41] (the dataset contains 10 million images categorized under 365 different categories) was used, and for testing it, MIT Indoor 67 dataset [42] (the dataset contains 15,620 images of indoor scenes categorized under 67 different categories) was used. The overall implementation was for recognizing the different scenes/places. The same task is also achieved using CNNs [43].

Further, there is an invention of capsule neural network (CapsNet) and spiking neural network (SNN). The results and the way of working of these neural networks suggest that they can act just like human brains. Geoffrey Hinton invented the capsule neural network. In CapsNet, capsules replace the neurons. They instead of using the process of rendering use the inverse graphics process of storing any data in memory just like humans. Therefore, if it stores an image seen with certain orientation and the same image when shown to it with some other orientation, then it easily recognizes it. It is because instead of taking the input and producing the output in scalar form, it takes the input and produces the output in vector form. Moreover, capsules of CapsNet unlike neurons not only perform weighting, sum and have an activation function, but also performs an affine transform which is used to store points, planes, and straight lines. One more interesting thing to notice in CapsNet is the dynamic routing to connect the capsules among them, which is another state of the art implemented to develop the neural network. The main reason behind the formation of capsules is because of the deficiencies found in CNNs. The major deficiencies in CNNs because of which the capsules came into existence are, the max-pooling layers that are supposed to increase the computation power and also decrease the size of the data. However, in doing so, they are losing the many information about the data. Also, the percept present at the end of CNNs, that contains all the information of the objects cannot form the correct orientation of a particular image [44-46].

SNN was first proposed in 1952 by Alan Lloyd Hodgkin and Andrew Huxley. The interesting part of this neural network is that the exchange of information among the neurons is accomplished using the chemicals in synaptic gaps instead of the spikes. There are many architectures proposed for SNN over the years. However, there is no specific architecture for it. Some of the different architectures proposed for SNNs are Hodgkin and Huxley model which is the first proposed SNNs, spike response model (SRM), Izhikevich neuron model and the most popular model termed as leaky integrated and fire (LIF) model. LIF is very popular because of its quality of adding a clear threshold along with charges across the leaky cell membrane that acquires the necessary instinctive properties from the input. The problem associated with SNN is that even today there is no specific algorithm developed for it to learn the different things. However, some of the learning rules developed for the SNNs are Unsupervised Learning by Spike Timing Dependent Plasticity (STDP), probabilistic approaches towards STDP and the use of supervised learning. The interesting thing about SNN is that it uses very low power. However, it has a very low accuracy rate [47-49].

9. Discussion and Conclusion

The earlier section demonstrated the different research and development work done with exceptional precision. Achievements of different goals using various techniques, methods, and systems demonstrated a perfection level of more than 99% in some cases.

MCDNN has proved itself best in recognizing the traffic signal. MCDNN is also used for image classification, object detection, character classification and so on. Moreover, in the field of traffic signal recognition, the use of a committee of the convolutional neural network (CNN) proved to have the same efficiency of recognizing various traffic signals with different shapes and sizes. Both techniques are multi-purpose because they are used in different tasks and have proven to be very accurate. However, there is an absence of comparison among these two by different tasks.

The system developed for visually impaired people using the 360^0 cameras has a precision of 92.8%, but it takes around 7 seconds for completing the whole process which is too long. This system indeed cannot be used by the visually impaired people because it takes a lot of time and before it will produce an output the person could have moved to a different scene. Therefore, the system does nothing to decrease any unwanted accident that the visually impaired person may meet.

Another novel system developed for the visually impaired people specifically for indoor navigation is an innovative system, but the system is a smart one, not an intelligent one. Moreover, for this, a person always needs to carry the smartphone which contains the required application. So, that makes it bit complicated and expensive. Also, the system demands the visually impaired person to operate the smartphone, so that cannot be considered practical.

AlphaGo program is basically for playing the game of Go. However, it proved itself to be the most intelligent program in the world of machine learning until now. AlphaGo program has also shown its mettle in clinical medicine, i.e., it will decide which medications to prescribe to a patient. Therefore, this program may prove to be one of the best in automatic navigation, object detection, and recognition because of its high accuracy and intelligence. So, implementing this technique for accomplishing different tasks may produce some interesting results.

In the field of scene recognition the work done or achieved is very less as compared to the work done in other sectors. Even after having so many state-of-the-art neural network architectures scene recognition is still one of the major problems. Scene recognition is one of the most important aspects of robotics and specifically in indoor assistive robotic systems. Therefore, the task of scene recognition using different neural networks requires proper direction as there are many challenges associated with it. One of the major challenges is that scene recognition on the external scenes have a higher accuracy rate as compared to indoor

scenes. Because according to studies and observations the accuracy rate of scene recognition on indoor scene drops drastically as compared to the external scenes. This problem persists and requires a proper solution.

Also, there are developments of various neural networks specifically for object detection and recognition. However, some of them have good accuracy but not the good speed, and some of them have appreciable speed but not the accuracy. Also, there are neural networks discussed in the earlier section which work almost like the human brain, however, still are implemented and tested on a very small scale. Therefore, it is necessary to implement state-of-the-art neural networks on a larger scale for different purposes so that their advantages and disadvantages over other neural networks can be classified and the logic behind having a proper neural network that emulates human brain can be developed.

Furthermore, it is very important to notice that there are so many aspects present in AI. However, the fusion of concepts done to date is very less. There is very fewer research works demonstrating the use of more than two concepts of AI. Either there is a neural network with fuzzy logic, or there is a neural network with GA. Therefore, this raises a very important question that whether in future a system with all concepts of AI will exist or not? And this question becomes more interesting because of the development of an ES called ROSS which has the traditional concept of ES along with the concept of deep learning. So the question now becomes that whether in future a system with the concepts of traditional ES, fuzzy logic, GA, PSO, fractional calculus, machine and deep learning, neural networks, etc. will exist or not? Because works on these kinds of systems are very less and if such works start happening through collaborative research then it may become possible to witness such systems, and such systems will become a benchmark in the history of humanity as they will have exceptional quality of performing several tasks like a human.

References

[1] N. Harris, "The Design and Development of Assistive Technology," *IEEE Potentials*, vol. 36, no. 1, pp. 24-28, 2017.

[2] S. Russell, P. Norvig, and A. Intelligence, "A modern approach Artificial Intelligence," *Prentice-Hall, Egnlewood Cliffs*, vol. 25, p. 27, 1995.

[3] V. Novák, I. Perfilieva, and J. Mockor, "Mathematical principles of fuzzy logic," *Springer Science & Business Media*, vol. 517, 2012.

[4] C. W. De Silva, "Intelligent control: fuzzy logic applications," CRC press, 2018.

[5] J. Sabatier, O. P. Agrawal, and J. A. T. Machado, "Advances in fractional calculus," *Springer*, vol. 4, no. 9, 2007.

[6] D. Baleanu, J. A. T. Machado, and A. C. J. Luo, "Fractional dynamics and control," *Springer Science & Business Media*, 2011.

[7] T. Li, G. Shao, W. Zuo, and S. Huang, "Genetic Algorithm for Building Optimization: State-of-the-Art Survey," *Proceedings of the 9th International Conference on Machine Learning and Computing*. ACM, pp. 205-210, 2017.

[8] G. M. Khan, "Evolutionary computation," *Evolution of Artificial Neural Development*: Springer, pp. 29-37, 2018.

[9] P. Jackson, "Introduction to expert systems," *Addison-Wesley Longman Publishing Co. Inc.*, 1998.

[10] J. McCarthy, "Some expert systems need common sense," *Annals of the New York Academy of Sciences*, vol. 426, no. 1, pp. 129-137, 1984.

[11] E. H. Shortliffe, R. Davis, S. G. Axline, B. G. Buchanan, C. C. Green, and S. N. Cohen, "Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system," *Computers and biomedical research*, vol. 8, no. 4, pp. 303-320, 1975.

[12] E. Shortliffe, "Computer-based medical consultations: MYCIN," *Elsevier*, 2012.

[13] S. Liberatore, "Your AI lawyer will see you now: IBM's ROSS becomes world's first artificially intelligent attorney," *MailOnline-news, sport, celebrity, science and health stories*, 2016.

[14] L. Luoren and L. Jinling, "Research of PID control algorithm based on neural network," *Energy Procedia*, vol. 13, pp. 6988-6993, 2011.
[15] S. Rajasekaran and G. A. V. Pai, "Neural networks, fuzzy logic and genetic algorithm: synthesis and applications (with cd)," *PHI Learning Pvt. Ltd.*, 2003.
[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
[17] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097-1105, 2012.
[19] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," *European conference on computer vision*. *Springer*, pp. 818-833, 2014.
[20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
[21] C. Szegedy, W. Liu, Y. Jia and P. Sermanet, "Going deeper with convolutions," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9, 2015.
[22] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016.
[23] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *European conference on computer vision*. *Springer*, pp. 630-645, 2016.
[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818-2826, 2016.
[25] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," In *AAAI*, vol. 4, p. 12, 2017.
[26] A. Géron, "Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems," *Sebastopol, CA: O*, 2017.
[27] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255-260, 2015.
[28] D. CireşAn, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column deep neural network for traffic sign classification," *Neural Networks*, vol. 32, pp. 333-338, 2012.
[29] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition," *Neural networks*, vol. 32, pp. 323-332, 2012.
[30] M. Ali, "360 View Camera Based Visual Assistive Technology for Contextual Scene Information," *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2135-2140, 2017.
[31] K. Chaccour and G. Badr, "Computer vision guidance system for indoor navigation of visually impaired people," *IEEE 8th International Conference on Intelligent Systems*, 2016, pp. 449-454, 2016.
[32] D. Silver *et al.*, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484-489, 2016.
[33] Z. Zhang, "When doctors meet with AlphaGo: potential application of machine learning to clinical medicine," *Annals of translational medicine*, vol. 4, no. 6, 2016.
[34] R. Girshick, "Fast r-cnn," *Proceedings of the IEEE international conference on computer vision*, pp. 1440-1448, 2015.
[35] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 6, pp. 1137-1149, 2017.
[36] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431-3440, 2015.
[37] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," *IEEE International Conference on Computer Vision*, pp. 2980-2988, 2017.
[38] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788, 2016.
[39] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," *arXiv preprint*, vol. 1612, 2017.
[40] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
[41] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba, "Places: A 10 million image database for scene recognition," *IEEE transactions on pattern analysis and machine intelligence*, 2017.
[42] A. Quattoni and A. Torralba, "Recognizing indoor scenes," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 413-420, 2009.

[43] L. Wang, S. Guo, W. Huang, Y. Xiong, and Y. Qiao, "Knowledge guided disambiguation for large-scale scene classification with multi-resolution CNNs," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 2055-2068, 2017.

[44] G. E. Hinton, A. Krizhevsky, and S. D. Wang, "Transforming auto-encoders," *International Conference on Artificial Neural Networks*. *Springer*, pp. 44-51, 2011.

[45] G. E. Hinton, S. Sabour, and N. Frosst, "Matrix capsules with EM routing," 2018.

[46] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," *Advances in Neural Information Processing Systems,* pp. 3856-3866, 2017.

[47] W. Maass, "Networks of spiking neurons: the third generation of neural network models," *Neural networks*, vol. 10, no. 9, pp. 1659-1671, 1997.

[48] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Transactions on neural networks*, vol. 14, no. 6, pp. 1569-1572, 2003.

[49] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, and A. S. Maida, "Deep Learning in Spiking Neural Networks," *arXiv preprint arXiv:1804.08150,* 2018.