2023 International Conference on Artificial Intelligence and Smart Communication (AISC) | 979-8-3503-2230-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/AISC56616.2023.10085166

A Thorough Review on Deep Learning Neural Network

Payal Chhabra
Research Scholar, Department of C.S.E
MM Engineering College, MMDU
Mullana-Ambala, Haryana, India
Payal49691@gmail.com

Dr. Sonali Goyal
Associate professor, Department of C.S.E
MM Engineering College, MMDU
Mullana-Ambala, Haryana, India
Sonaliguglani21@gmail.com

***Abstract*-Deep Neural Network Topologies are used by subfield of machine learning called "deep learning" that are similar but different to handle a variety of challenges in domains including bioinformatics, computer vision and, among others. Recent research on deep learning has grown significantly across several applications. Deep learning technique produces state-of-the-art results by using numerous layers of features or data representations. Deep learning is essentially the application of neural networks with multiple hidden layers of neurons. In particular, this review paper firstly aims to offer a more thorough overview of the most fundamental deep learning components, further shows how deep learning techniques outperformed well-known ML techniques and then outlines how to deploy and build deep learning model. Secondly, Convolutional neural networks (CNN), most common used deep learning networks, are then introduced, along with a description of how they have implemented with matrix representation. Thirdly, we concentrate on application domain of deep learning, with an emphasis on its use in object detection. Finally, Future study directions are provided after the conclusion of the publication to assist scholars in understanding the research gaps and findings.**

Keywords: Deep Learning, Machine Learning, Convolutional neural networks, forward propagation, Object Detection

I. INTRODUCTION

A branch of artificial intelligence known as "deep learning "that replicates how people learn specific types of expertise. A key component of records technology, which includes statistics and predictive modeling, is deep learning [1]. Information scientists who must gather, analyze, and interpret large amounts of statistics will greatly benefit from deep learning because it will speed up and simplify the process. Deep studying can be used as a basic form of automation for predictive analytics. The learning algorithms in conventional systems are linear, whereas the learning algorithms in deep mastering systems are stacked in a hierarchy of increasing complexity and abstraction.

Deep learning instructs computers to learn by doing what comes naturally to humans. Driverless cars use this technology to actually recognize stop signs and other signals [2]. It is essential for voice control on consumer electronics like hands-free speakers, smart phones etc. Recently, it has attracted a lot of attention for good reason. It produces outcomes that were previously unattainable. Using deep learning, model learns to perform classifications automatically from text, sound, as well as images. Deep learning models can occasionally do better than humans and reach the highest levels of accuracy. Models are trained using neural network models and huge collection of labeled data. Despite being first theorized in the 1980s, deep learning has only recently started to be useful for two main reasons:

- Deep learning required large labeled data. For instance, creating driverless car requires very large dataset of images and video.
- Deep learning requires a substantial amount of processing power.

Deep learning continually examines data by predetermined logical structure in an effort to reach conclusions that are comparable to those reached by humans. Deep learning accomplishes these using neural networks, a multi-layered structure of algorithms.

A. *Process of Deep Learning*

In deep learning majority of learning techniques use neural network architectures in which number of hidden layers is typically indicated by the term "deep" [3]. While deep networks can have as many hidden layers as they like, traditionally, neural networks include two or three layers. To train these types of models, huge amounts of labeled input and neural network (automatically extract features) are used.

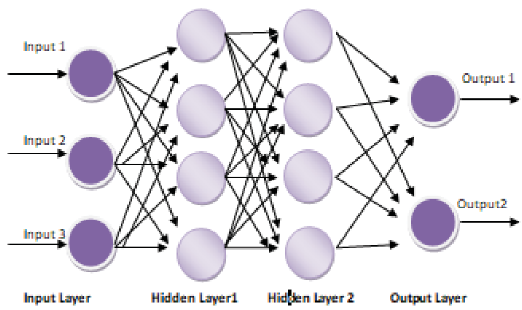

Fig. 1. Architecture of Deep Neural Network

B. *How Machine Learning is differ from Deep Learning*

Machine learning manually extracting pertinent features from images and then model which classifies objects from image is developed using these features. Using deep learning methodology, pertinent features are automatically retrieved from photos [4]. End-to-end learning is another function it does. In this method, a network is given raw data and a goal to accomplish such classification, and then automatically learns or known how to execute it.

1) No feature extraction: First, deep learning has several advantages over machine learning, including the redundant nature of so-called feature extraction [5]. We used machine learning techniques like SVM, logistic regression long before we started using deep learning. Flat algorithms are another name for these algorithms. Typically, these techniques cannot be used directly on the raw data. The term "flat" is used here. It is necessary to do a feature extraction preprocessing stage. These traditional machine learning algorithms can use the representation of the given raw data created by feature extraction to complete a task. We can now, for instance, categories the data into different classes. A thorough understanding of the problem area is usually required for the difficult process of extracting features. Preprocessing layer needs to be enhanced throughout numerous iterations in order to produce the best outcomes. Artificial neural networks based deep learning does not have the feature extraction stage. These layers have the ability to independently and directly learn an implicit representation of raw data. Here's how it works: An artificial neural network's numerous layers can be used to represent the raw data that is increasingly compressed and abstract is created. The result is then generated using this compressed version of the input data. The outcome could be the categorization of the input data into various classes.

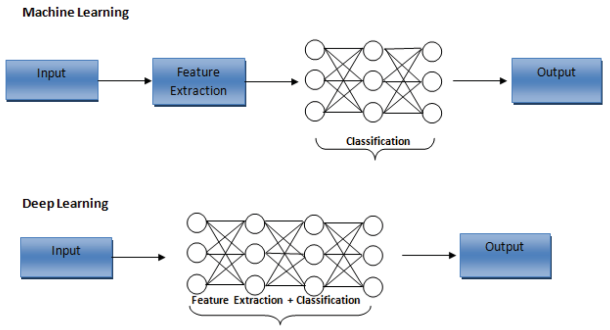

Fig. 2. Deep learning v/s Machine Learning

2) *The era of big data:* Deep learning's second major benefit, which is also a key factor in understanding why it is growing in popularity, is that it is driven by enormous amounts of data. Huge opportunities for new deep learning innovations will arise in the big data era. In contrast to SVM and naive Bayes machine learning models, Deep learning models often see their accuracy rise with the volume of training data, unlike traditional learning algorithms, which plateau after a certain point [6].

C. Steps to build neural network

1) *Identification of appropriate deep learning function:* Deep Learning function includes computer vision, classification, localization, detection [7].

- Computer Vision: A branch of artificial intelligence that allows humans to teach computers how to comprehend visual data from pictures and videos by using models and algorithms [8].
- Image Classification: Uses artificial intelligence to detect and label the images [9]. User given features are used to classify these images.
- Object Localization: Object localization includes identifying distinct things in a visual and putting a boundary around them—typically a box—to categories them. A single object's location can be determined using image localization.
- Object Detection: Both of these methods are used, and the objects are classified before being divided into smaller groups and given labels based on their characteristics [10]. By identifying and classifying one or more items in an image, object detection combines the two.

2) *Pick a framework:* PyTorch and Tensor- flow are the tools for training and testing the neural network which are typically included in a framework or toolset used to construct a neural network.

3) *Assemble Data sets for the neural network:* Depending on the kind of data a neural network will analyze, a certain number of images must be collected for training. Typically, a collection of training images is needed for every feature and every grade of that feature that the neural network must evaluate. The neural network can learn to evaluate those categories more precisely with more images that are presented for each category.

Image Datasets – MNIST, MS –COCO, ImageNet

Natural Language Processing- IMDB Reviews, Sentiment140, the Wikipedia Corpus

Audio/Speech Datasets- Free Spoken Digit Dataset, Free Music Archive (FMA), VoxCeleb

4) *Train and validate:* To ensure accuracy by separating training from test data, it is possible to prevent a neural network from unintentionally learning from subsequent evaluation data. This process can be quickened by applying transfer learning or utilizing a network that has already been trained for another task [11]. For example, a neural network that has been trained for feature extraction may just require a new set of photos to recognize a new feature. Pre-trained networks are offered for free by frameworks like Caffe2 and TensorFlow.

5) *Deploy and test on new data:* The final phase involves deploying a trained neural network on the chosen hardware in order to evaluate performance and gather data in the real world. The use of the cloud enables speedy scaling up, the deployment of modifications across several sites, and significant cost savings on hardware. However, problems with internet connections might result in catastrophic failures, and cloud deployment has a higher latency than edge deployment.

II. CONVOLUTIONAL NEURAL NETWORK

Most commonly employed in Deep Learning is neural network. It employs 2D Convolutional layers and combines input data with learned features, making it an excellent architecture for processing 2D data, including images. There is no need to know the features for classification images because CNNs do manual feature extraction. The direct extraction of features is just how CNN functions [12, 13, 14]. As the network learns from an image, crucial features are found rather than being pre-trained; this is the work of neural network. For computer vision-based applications like object categorization, deep learning is accurate, thanks to this automated feature extraction. Using dozens or perhaps hundreds of hidden layers, CNNs can learn to recognize various features from an image.

A. CNN Implementation

Convolutional Neural Network specifically created for images and videos. It learns the features of the images it receives as inputs; extracts data from them, then categorizes the images using the learned features [15, 16]. This program takes its ideas from Visual Cortex functions of human brain. Processing of visual data from the outside world is carried out by visual cortex. It comprises several numbers of layers and each layer functions independently, extracting different information from images or other visuals. When information from each and every layer has been combined, the picture is then evaluated and labeled. Similar to this, CNN uses a variety of filters that each collect data from the image, such as edges and various forms (vertical, horizontal, and round), which are then combined to identify the image.

Artificial Neural Networks cannot be used for the same purpose because of some limitations [17,18]:

- For an ANN model to be trained, large and diverse image sets require too much computing.
- ANN cannot retrieve entire relevant information from an image, but a CNN model can collect a pixel's value that is influenced by nearby pixel's value in image (spatial dependencies).
- The placement of the object in the image is important to ANN; if the location or site of the same object varies, the classification will be incorrect.

B. *Components of CNN*

The two stages of the CNN model's operation are feature extraction and classification. The process of extracting information and features from images is called feature extraction, and once that process is complete, the images are passed on to the classification phase, where they are categorized according to the problem's target variable

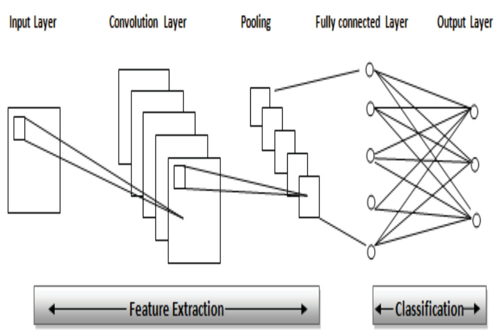

Fig. 3. CNN Layered Architecture

1) *Input layer:* Input image, as the name implies, can be either RGB or Grayscale. Each image is composed of pixels with values between 0 and 255. Before sending the data to the model, it need to be normalized, or from range (0, 1). An example of an input image with three channels—RGB and pixel values—and a size of 3*3 is shown below.

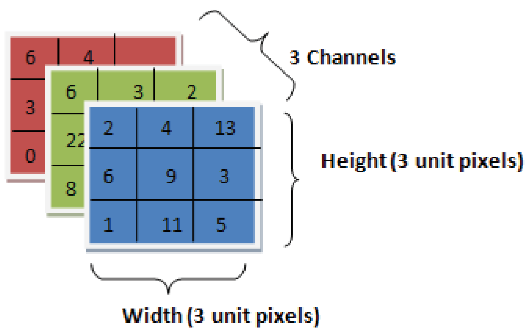

Fig. 4. Input Layer

2) *Convolution Layer:* The filter used on input image during this layer in order to extract or detect its characteristics. Multiple filtering operations are performed on the image to produce feature map that categorize an input image. Let's use an illustration of 2D input image with normalized pixels for better understand this.

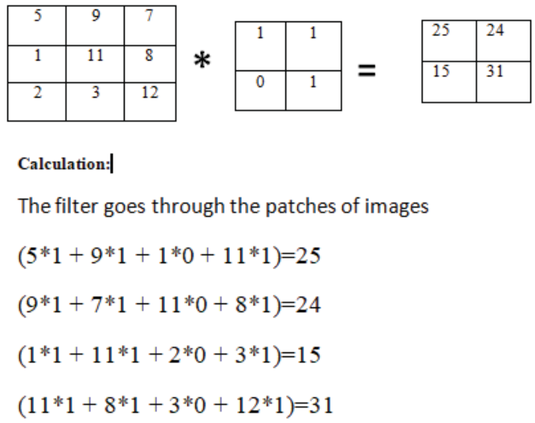

Fig. 5. Filter on Convolutional Layer

In the example above Fig. 5, we applied a (2, 2) filter on an input image of size (3, 3) in order to identify specific features. While we only used one filter, in reality, many of these kinds of filters are used to extract data from images. When the filter is applied to the image, we obtain a feature map with the coordinates (2, 2) that contains some data of input image. The entire image is covered by the filter, and the result is final Feature Map. Once feature map has been obtained, nonlinearity is added by applying an activation function to it.

a) *Stride:* The stride of 1 moves the weight matrix by a single pixel at a time, and as the stride value increases, the size of image keeps getting smaller. The feature map's size is smaller than the size of the corresponding image, and it gets smaller as the stride value increases. For more complicated inputs or filter dimensions, use the following formula to get the output's dimension:

Dimension of image = (x, y)
Dimension of filter = (f, f)
Dimension of output = ((x-f+1), (x-f+1))

3) *Pooling Layer:* This layer decreases feature map's size, helping in the preservation of the input image's significant information or features while speeding up computation. With pooling, reduced resolution version of the input is produced which still includes the original image's significant or major components. It is of two types:

a) *Max Pooling:* The brighter pixels in the image are chosen via max pooling. It is helpful when we just care about the image's lighter pixels and the image's dark background. After performing Pooling, the size of the feature map has decreased [19, 20]. Maximum value is taken from each highlighted area, producing a 2*2 size of input image.

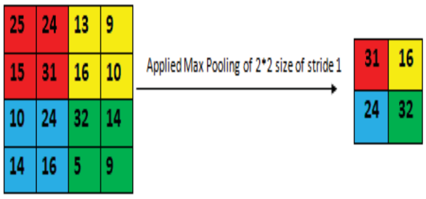

Fig. 6. Max Pooling

b) *Average Pooling:* This preserves a great deal of data about the "less significant" components of a block or pool [19, 20]. Average Pooling combines them, as opposed to Max Pooling, which simply discards them by selecting the highest value. This is helpful in many circumstances where such information is helpful.

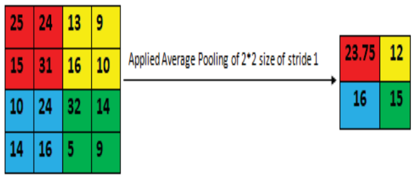

Fig. 7. Average Pooling

4) *Fully Connected Layer:* After processing the Feature Extraction; the next step is Classification. The input image is classified into a label using this layer only [21]. Then both Convolutional and pooling information connected to output layer through this layer, which ultimately labels input according to the classification criteria.

To this point, the convolution layer has taken the features from the data and sent them to connected layer to produce results. The convolution layer produces a 2D matrix as an output, whereas the fully connected layer only processes 1D data. Consequently, the 2D data were transformed to 1D format first.

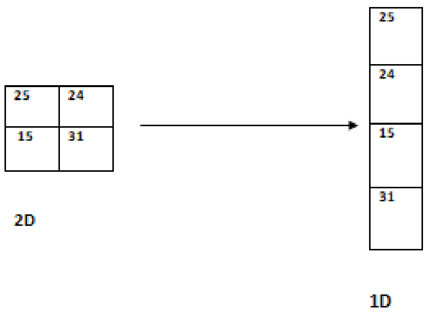

Fig. 7. Conversion from 2D to 1D

The data is delivered to the fully connected layer after it has been transformed to 1D. Every single value is handled as a distinct characteristic that makes up an image. It executes two operations e.g. linear and non-linear transformation

a) Linear transformation Equation:

$$Z = W.X + b \tag{1}$$

Where, W is matrix number of weight, X is input and b is constant (bias).

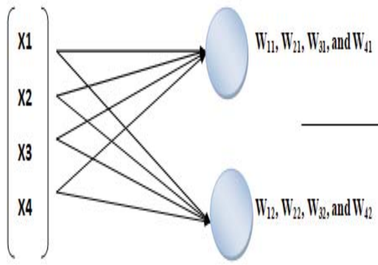

Fig. 8. Linear Transformation

(m1, m2) size matrix, where m1 = number of features and m2 = number of neurons in layers. Like in above example 4 features are taken so, m=4 and if it has two neurons then shape of weight matrix will be (4, 2)

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} \quad W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \\ W_{31} & W_{32} \\ W_{41} & W_{42} \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Input Data | Randomly Initialized Weight Matrix | Randomly Initialized bias Matrix

Fig. 9. Matrix Representation

Let put weight and bias matrix in the linear equation transformation:

$$\mathbf{Z = W.X + b}$$

$$Z = \begin{bmatrix} W_{11} & W_{21} & W_{31} & W_{41} \\ W_{12} & W_{22} & W_{32} & W_{42} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$Z_{2x2} = \begin{bmatrix} W_{11}X_1 + W_{21}X_2 + W_{31}X_3 + W_{41}X_4 \\ W_{12}X_1 + W_{22}X_2 + W_{32}X_3 + W_{42}X_4 \end{bmatrix}$$

Fig. 10. Matrix representation of Linear Transformation

b) *Non-Linear transformation:* Because complicated relationships cannot be handled by linear transformation alone, non-linearity is added to the data by including a new element called an activation fnction in each layer of the neural network [22, 23].

Sigmoid activation for the binary classification problem,

$$S(x) = 1/\,(1+e^{-x)} \tag{2}$$

Where, S(x) = sigmoid lies between range (0, 1)

e= is Euler's function.

C. Forward propagation:

1. Load an input variable(x) with images

2. Define the filter matrix (f)

Z1=x*f

3. On the result, apply sigmoid function

A=sigmoid (Z1)

4. Define the weight and bias matrices, and then translate the data linearly by apply linear transformation on those values.

Z2=W.A + b

5. Again use the sigmoid function on the results of the linear transformation

O=sigmoid (Z2)

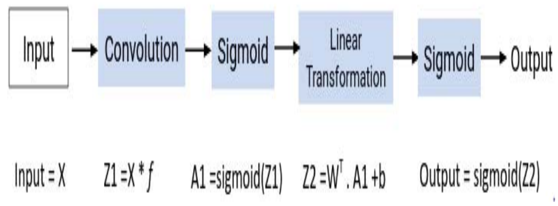

Fig 11. Forward Propagation

II. LITERATURE REVIEW

Asifullah khan et al. discussed about seven fundamental taxonomies of newly implemented CNN based on spatial exploitation; channel boosting, depth, multi-path, feature-map exploitation, attentiveness and breadth [24]. Furthermore, Basic components, applications and current difficulties of CNN are mentioned.

Anamika Dhillon et al. present a comprehensive analysis of numerous deep architectures and models, focusing on the unique properties of each model [25]. First, it presented the working of CNN architectures and their individual components, followed by many CNN models, beginning with the traditional LeNet model and including AlexNet, GoogleNet, VGGNet and many more. This primarily concentrates on application of deep learning to detect animals, small arms, and human.

Anjeel Upreti presented about how CNN is useful for challenging tasks including classification, object detection, recommendation systems, and natural language processing [26]. Its extensive network effectively completes the mission. This publication presented a thorough explanation of the model structure that researchers interested in using the model might use as a reference point. To give researchers interested in applying CNN models in their work a reference point or baseline, the fundamentals of CNN models and their application are discussed in this paper.

Tejal Tiwary et al. proposed the automatic creation of image captions using development of deep learning-based interaction with people is advantageous to blind users [27]. The extended Convolutional atom neural network is useful tool to creating image captions. CNN is used to create pre-trained captions, and long short-term memory implements a reverse search to choose caption that is appropriate for image. Suggested paradigm has a drawback having no such hardware platforms, as smart phones and glasses are integrated. In the future, it can be expanded to examine how well it performs on huge datasets for MSCOCO or Flickr30k to create accurate captions

Ruchita Tatkare et al have introduced a new Object Detection method which is done by taking an image as an input and then goes through Mask-RCNN algorithm for detection [28]. Using a system which is stationary by using Smartphone, we can overcome this problem. Smartphone is portable so it comes easy to carry.

Shanthi K G et al. have suggested proposing a system that consists of an audio amplifier module, tensor flow, memory card, Raspberry Pi 3 processor, and pi camera [29]. To detect objects, the Solid-State Drive with Mobile net SSD was selected. However, the usage of Amazon's online cloud service, which saves all the programs and data, can minimize the price of the Raspberry Pi board.

Sangeeta Mahapatra et al suggested to uses to help the blind find and recognize objects in their environment; you only look once v3 algorithm had been trained using Open CV, TensorFlow, and the COCO Dataset [30]. To create audio from the specified object, use Google Text to Speech. The proposed system also leads Yolov3 can analyze video and image data to find 80 different items. It is more accurate. It functions as both an object detector and an object locator. YOLOv3 operates at 28.2 maps in 22ms at 320x320. There are now 3 distinct detection scales.

III. APPLICATIONS

There are wide variety of deep learning applications including automated driving and medical equipment.

- Automated Driving: Deep learning is being used by automotive experts to automatically recognize items like stop signs and traffic lights [31]. In order to identify pedestrians and prevent accidents, deep learning is also applied.
- Aerospace and defense: Deep learning recognizes objects from satellites to identify points of interest and to categories troops' operating environments into safe and unsafe areas.
- Medical Research: To automatically identify cancer cells, researchers studying cancer are using deep learning [32]. A multidimensional data collection produced by sophisticated microscope created by UCLA research teams was utilized to precisely train a deep learning application to identify cancer cells.
- Automation in Industry: By automatically determining when individuals or objects are too close to heavy machinery, deep learning is assisting in enhancing worker safety around such equipment.
- Electronics: Automated speech and hearing translation uses deep learning [33]. Deep learning software, for instance, is used to power voice-activated home assistance systems that remember your preferences.

A. Application of deep learning in object detection

Object is nothing but an element that can be visually represented. There is not much variation in an object's physical characteristics. For an object to be recognized and distinguished, it must be semi-rigid.

Deep learning-based object detection offers a quick and precise way to predict where an object will appear in an image. The object detector automatically learns the image features needed for detection tasks using deep learning. Deep learning-based object detection is teachable machine or computer technology that finds instances of particular classes of objects in pictures or videos [34, 35]. Object detection identifies those objects that may be present in an image or video stream and provide details such as the position. It might be able to recognize more than one object in the given photo in particular with respect to just one.

Identifying, locating, and detecting of many visual occurrences of things in a video or image are made easier with the aid of object detection techniques [36]. Compared to just basic object classification, it provides a much improved comprehension of the thing as a whole. The number of instances of distinct objects can be counted using this technique, and their precise locations can also be marked along with labeling. We can now use this process for real-time use cases thanks to a significant improvement in performance over time. It offers a complete answer to the question, "What object is where and how much of it is there?"

Humans are able to accurately detect different objects that are in front of us and identify each one of them. We can count and recognize many different objects with little effort. Large data sets are now readily available as a result of recent technological advancements, allowing computers to perform the same classification and detection tasks as humans.

IV. CONCLUSION

An overview of deep learning was provided in this paper, including Convolutional neural networks. These models can be thought of as fundamental deep learning architectures currently. Additionally, we covered related ideas like activation function and linear and non-linear transformation that are necessary for technical knowledge of these models. The key to being prepared for future advancements is therefore having a fundamental understanding of these components. Finally, we need to take into consideration the fact object detection systems is required for nano-robots or for robots. These object detection will also need to learn new classes when it comes.

Some basic developments in deep learning and their numerous applications have been covered in this paper. Finally, additional presentations of deep learning applications are made. Deep learning can occasionally achieve unexpected and better performances in many areas, such as image processing and object detection, which are highly challenging for persons with visual impairments to notice. This is because there are so many problems that are being solved daily. Therefore, a development in one area may lead to a game-changing idea in a different area. Deep learning is gaining popularity quickly; new applications or discoveries are being made daily. Following are a few topics of active research that, to the best of our knowledge, will continue to attract interest in the near future.

REFERENCES

[1] S. Pouyanfar,‘‘A survey on deep learning: algorithms, techniques, and applications,’’ ACM Comput. Surv., vol. 51, no. 5, p. 92, 2018.
[2] Chirag Sharma,”Self driving car using deep learning technique”, International Journal of Engineering and Technical Research V9(06), 2020
[3] Iqbal Sarker, ”Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions”, SN Computer Science vol2, 2021
[4] Alejandro Baldominos, Alejandro Cervantes,“A comparison of machine learning and deep learning techniques for activity recognition using mobile devices”, Sensors 19(3):521, DOI:10.3390/s19030521, 2019
[5] Hind Alaskar, Tanzila Saba, “Machine learning and deep learning: a comparative review”, In book: Proceedings of Integrated Intelligence Enable Networks and Computing, IIENC 2020, 2021, DOI: 10.1007/978-981-33-6307-6_15
[6] Alom MZ, Taha TM, Yakopcic, Westberg, Sidike, Nasrin MS, Hasan , Van Essen , Awwal, Asari, “A state-of-the-art survey on deep learning theory and architectures”, Electronics. 2019;8(3):292.
[7] Fatima Zahra, Hamza Bouftaih, “Simultaneous object detection and localization using convolutional neural networks”, DOI:10.1109/ISACV.2018.8354045 Conference: ISCV'18, 2018.
[8] Bengio, Lecun, Hinton, “Deep learning for AI”, Communications of the ACM, July 2021, Vol. 64 No. 7, Pages 58-65 10.1145/3448250.
[9] Qing Lv,1Suzhen Zhang,Yuechun Wang,” Deep learning model of image classification using machine learning”, ID 3351256, https://doi.org/10.1155/2022/3351256,2022.
[10] M.Thariq, D. Saidulu, A. Manikandan, P. Naresh, ” Object detection and recognition in real time using deep learning for visually impaired people” , International Journal of Electrical and Electronics Research (IJEER),2022.
[11] Asmaul Hosna, Ethel Merry, Jigmey Gyalmo, Zulfikar Alom, Zeyar Aung & Mohammad Abdul,” Transfer learning: a friendly introduction” Journal of Big Data,2022.
[12] R. Vargas, A. Mosavi, and R. Ruiz, ‘‘Deep learning: A review,’’ in Proc. Adv. Intell. Syst. Comput., 2017, pp. 1–11.
[13] Laith Alzubaidi, Jinglan Zhang, Amjad Humaidi, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions”, Journal of Big Data volume 8, 2021.
[14] Dhillon A, Verma GK.,”Convolutional neural network: a review of models, methodologies and applications to object detection”, Prog Artif Intell. 2020;9(2):85–112.
[15] Q. Zhang, M. Zhang, T. Chen, Z. Sun, Y. Ma, and B. Yu, “Recent advances in convolutional neural network acceleration,” Neurocomputing, vol. 323, pp. 37–51, 2019.
[16] A. Gülcü, Z. KUş, “Hyper-parameter selection in convolutional neural networks using microcanonical optimization algorithm,” IEEE Access, vol. 8, pp. 52528–52540, 2020.
[17] Mehmood Hasan1,Saleem Ullah, Muhammad Jaleed Khan, Khurram Khurshid, ”Comparative analysis of svm, ann and cnn for classifying vegetation species using hyperspectral thermal infrared data”, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W13, 2019
[18] Muhammad Asif, Norhalina Senan, Fazli, Muhammad Aamir, Ali Samad, Mukhtaj Khan,” Comparative analysis of recent architecture of convolutional neural network, ID-7313612, https://doi.org/10.1155/2022/7313612, 2022.
[19] Hossein Gholamalinejad, Hossein Khosravi,“ Pooling methods in deep neural networks, a review”, September 2020.
[20] Afia Zafar, Muhammad Aamir, Nazri, Ali Arshad, Saman, Abdulrahma , Ashit Kumar Dutta and Sultan Almotairi,” A comparison of pooling methods for convolutional neural networks”, Appl. Sci. 2022, 12, 8643. https:// doi.org/10.3390/app12178643.
[21] Sh Shabbeer Basha, Shiv Ram, Viswanath Pulabaigari, Snehasis Mukherjee, “Impact of fully connected layers on performance of convolutional neural networks for image classification”, Neuro computing 378, DOI:10.1016/j.neucom.2019.10.008,2019.
[22] Siddharth Sharma, Simone Sharma, Anidhya Athaiya,” Activation functions in neural networks”, International Journal of Engineering Applied Sciences and Technology, 2020.
[23] Murilo Gustineli,” A survey on recently proposed activation functions for deep learning”, arXiv: 2204.02921v2 [cs.LG] 7, Apr 2022.
[24] Asifullah Khan, Anabia Sohail, Umme Zahoora & Aqsa Saeed Qureshi ,” A survey of the recent architectures of deep convolutional neural networks”, Artificial Intelligence Review volume 53, pages5455–5516,2020.

[25] Anamika Dhillon & Gyanendra Verma," Convolutional neural network: a review of models, methodologies and applications to object detection", Progress in Artificial Intelligence volume 9, pages85–112 ,2020.
[26] Anjeel Upreti, "Convolutional neural network (cnn). a comprehensive overview", DOI:10.20944/preprints2022 08 .0313.v3, 2022.
[27] Tejal Tiwary1 & Rajendra Prasad," An accurate generation of image captions for blind people using extended convolutional atom neural network" in Springer Science, 2022.
[28] Mayuresh Banne, Rahul Vhatkar, Ruchita Tatkare "Object detection and translation for blind people using deep learning", International Research Journal of Engineering and Technology (IRJET),2020.
[29] Shanthi K "Smart vision using machine learning for blind" in International Journal of Advanced Science and Technology,2020.
[30] Sangeeta Mahapatra , Neha Chugwani, Pooja Chuathmal , Arti Tekade," Real Time object detection using yolo v3 tiny with voice feedback for visually impaired" ,International Journal for Research in Applied Science & Engineering Technology (IJRASET),2021.
[31] Darsh Parekh, Nishi Podd , Aakash Rajpurka, Manisha Chahal, Neeraj Kuma, Gyanendra Joshi and Woong Cho," A review on autonomous vehicles: progress, methods and challenges", https://doi.org/ 10.3390/electronics11142162 2022.
[32] Wenjin Yu, Yangyang Liu, "Deep learning-based classification of cancer cell in leptomeningeal metastasis on cytomorphologic features of cerebrospinal fluid", Sec. Cancer Imaging and Image-directed Interventions, 2022.
[33] Rushiraj Chavan,Vipul Jadhav,Yogesh Tekam,Ashutosh Pattanayak," Speech-to-speech translation using deep learning based models and cloud services",GIS science Journal,2022.
[34] Abdul Muhsin, Farah F., Bashra Kadhim ," Online blind assistive system using object recognition" ,in International Research Journal of Innovations in Engineering and Technology (IRJIET),2019.
[35] Nourhan , Anwar Awad and Talal Bonny "Smart assistant for blind and visually impaired people" in ICAAI ,2019.
[36] Meian Li, Haojie Zhu, Hao Chen, Lixia Xue, "Research on object detection algorithm based on deep learning". Journal of Physics Conference Series, DOI:10.1088/1742-6596/1995/1/012046, 2021.