ARCHITECTURAL PARADIGM OF DEEP LEARNING

Muhammad Arslan Ilyas[2], Umar Farooq[1]

ABSTRACT

Deep learning has been trendy and intriguing in recent years in machine learning. DL is the most accurate, supervised, time, and cost-effective way in terms of ML. There is no end to the amount of knowledge you may get with deep learning. Useful in a wide range of demanding situations, it covers a wide range of processes and structures. The approach categorizes the process of mastering the art of illustration. Security is one area where deep learning techniques have made significant progress. One of the most successful methods for uncovering complex patterns in large datasets is backpropagation. Biomedical image classification, object identification, cancer diagnosis, and many other applications of deep learning are the most commonly utilized domains of deep learning. Various aspects of deep learning, including their fundamental and advanced structures and methodologies as well as their motivational and other characteristics, are discussed in this work. Additionally, the article explores the key contrasts between classical ML and DL and the most pressing future challenges. The primary goal of this paper is to present an in-detail analysis of the significant applications of deep learning in several areas, including a look at the techniques and structures employed and the impact of each in practice.

INTRODUCTION

Using artificial intelligence, machines may learn new ideas and concepts on their own, without having to be explicitly programmed. Observations are the first step toward improved future outcomes and judgments being drawn from the data. The term "deep learning" means a group of ML algorithms that utilizes numerous nonlinear alterations to represent high-level notions gleaned from data. In the training phase, large datasets must be labeled and matching features found, but in the inferring phase, inferences are drawn and unknown datasets are labeled based on previous information. To help the system to understand and perform complicated sensory tasks with a high level of accuracy, deep learning is used. In the context of machine learning, deep learning is a technique that uses nonlinear processing units to extract and manipulate information from a huge volume of data. The outputs from the previous layer are sent into the subsequent levels. [1]

Figure1: *Traditional Machine Learning Architectural Pipeline*

Figure2: *Deep Learning Workflow Pipeline*

Deep learning uses a large quantity of data to define and classify patterns and combinations of attributes that may be used in a wide variety of contexts, including in medical diagnostics, financial markets, and more. The data that uses particular algorithms, numerous extensive features, and no human interaction is a significant feature of deep learning. For the sake of analyzing large datasets and eliminating spam communications, Facebook also developed Deep Text. The deep learning methodology is based on the following important factors:

- Nonlinear processing with various levels
- Supervised learning and unsupervised

As an instance of a hierarchical non-linear processing technique, the use of numerous layers are used, each receiving and passing on findings from the preceding layer. Hierarchical layers are used to control the significance of data. Unsupervised and supervised learning can be linked by using the category target label. The existence of this feature indicates a supervised system, while the absence of this feature demonstrates an unsupervised system. Soniya et al.[2] investigated the current state of deep learning methodologies, models, structure, and limits. The researchers looked into various approaches to training, optimization, and fine-tuning. They also emphasized the use of huge amounts of data and their application with deep learning. Deep learning was also brought up as a point of contention.

FUNDAMENTAL ARCHITECTURES OF DNN

Deep neural networks, recurrent neural networks, and belief networks are all terms used to describe deep learning architectures. Many layers are buried between the input and output layers of an Artificial Neural Network with a variety of topologies. It is possible to simulate complicated and non-linear interactions using a deep neural network and develop models that treat the item as a layered arrangement of primitives. Data travels from the input layer to the output layer in feedforward networks without any looping. It is possible to implement the notion of deep learning in a variety of ways.[1]

Year	Deep Learning Architecture
1990 - 1995	*RNN (Recurrent Neural Network)*
1995 - 2000	*LSTM (Long Short-term memory, CNN)*
2000 - 2005	*LSTM (Long Short-term memory, CNN)*
2005 - 2010	*DBN (Deep Belief Network)*
2010 - 2017	*DSN (Deep Stacked Network, Gated Recurrent Unit)*
2014 - 2020	*GAN (Generative Adversarial Network)*

In this paper, we will discuss seven basics architectures of deep learning, these are:

1. Recurrent Neural Network (RNN)
2. Convolutional Neural Network (CNN)
3. Restricted Boltzmann Machine (RBM)
4. Deep Stacking Network (DSN)
5. Long Short-Term Memory (LSTM)/Gated Recurrent Unit Network (GRU)
6. Auto-Encoder (AE)
7. Generative Adversarial Network (GAN)

RECURRENT NEURAL NETWORK

RNNs are a well-known and widely used technique in the field of deep learning [3,4,5]. There are several applications for RNNs in voice and natural language processing [6,7]. Recurrent neural networks (RNN) differ from regular networks by employing sequential information in their networks. Useful information may be gained from the data sequence's structure because of this feature. As an example, one needs to understand the statement's context to understand the original word in a phrase. RNN may be regarded as an STM unit, with the input layer (x), output (y), and state (hidden) layers.

$$h_t = \tanh\ h\,(W_{hh}.\,h_{initial} + W_{xh}.\,X_t)$$

$$h_{t+1} = \tanh h\,(W_{hh}.\,h_t + W_{xh}.\,X_{t+1})$$

$$h_{t+2} = \tanh\ h\,(W_{hh}.\,h_{t+1} + W_{xh}.\,X_{t+2})$$

Final Output:

$$y = W_{hy}.\,h_{t+2}$$

Hidden-to-hidden, Hidden-to-output, and Hidden-to-input are three forms of deep RNN algorithms invented by Pascanu et al.[8]. These three approaches are used to create a deep RNN, which has the advantages of a deeper RNN while reducing the learning complexity.

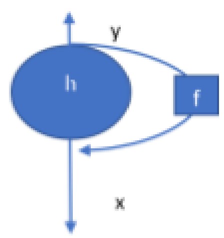

Figure3: *Pascanu defined RNN Algorithms structural parameters*

However, RNN's vulnerability to explosive gradients and disappearing difficulties is one of the key shortcomings of this technique[9]. During the training phase, replication of multiple big or tiny derivatives may lead to the slopes rapidly bursting or collapsing. As new inputs come in, the system forgets about the old ones, so susceptibility decreases over time. LSTM[10] can also be utilized to tackle these issues.

CONVOLUTIONAL NEURAL NETWORK

Deep learning algorithms like CNN are widely employed [11, 12, 13, 14, 15, 16] with no need for human involvement to identify crucial parts. CNN has a big edge over its predecessors. Computer vision[18], speech processing[19], and image recognition[20] are only a few of the areas where CNNs are used. CNNs, like regular neural networks, were inspired by actual brain cells. The occipital system in the cat's brain is represented by a complicated network of cells, and CNN [21] matches this arrangement.[22]

Goodfellow and his colleagues The training step is simplified while the network speed is increased using a minimal number of variables. It is just like the visual cortex cells. In an interesting twist, these cells don't recognize the whole image. Only a small part of it is identified by them. In other words, they "spatially extract" the input's local correlation and employ it as though it were a filter.

$$h^k = f(W^k * x + b^k)$$

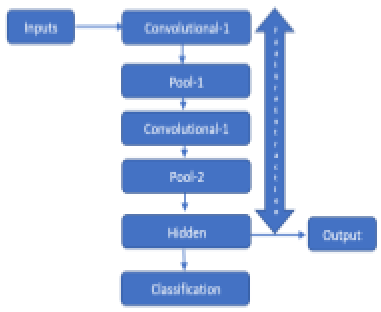

Figure4: *Convolutional Neural Network Architecture Diagram*

RESTRICTED BOLTZMANN MACHINE

To illustrate a hidden unit, a visible layer, and an unsupervised equal link between those levels, we use the RBM (Restricted Boltzmann Machine). In RBM, the hidden units are not connected to the input in any way. There are many hidden layers in the deep belief system, and it employs a novel training method. Stacks of restricted Boltzmann machines, or RBMs, are the building blocks of every connected layer. An abstract notion of this data is reflected in the hidden layer, which is conveyed by the input layer as sensory data. Classifying networks is the primary objective of the output layer. Early training and fine-tuning are two separate stages of training. In unsupervised pre-training, RBM can reproduce its input starting with the first hidden unit. The Second is similar to the first one in that the first hidden units are the source and visible layer, while the RBM is operating on the first hidden layer's outputs. Thus, each layer is pre-trained in its unique way. After the pre-training, the supervised fine-tuning process begins. [1]

Gibbs Sampling

$$p(h) = \frac{1}{1+e^{(-(a_i + w_i h_j))}} = \tilde{O}(a_i + \sum h_j w_{ij})$$

$$p(v) = \frac{1}{1+e^{(-(b_i + w_j v_i))}} = \tilde{O}(b_j + \sum v_j w_{ij})$$

Divergence:

$$p(h) = \frac{1}{1+e^{(-(a_i + w_i h_j))}} = \tilde{O}(a_i + \sum h_j w_{ij})$$

$$W_{new} + W_{old} + \Delta W$$

DEEP STACKING NETWORKS

The terms "Deep Convex Networks" (DCN) and "Deep Stacking Networks" (DSN) are interchangeable. DSN is different from other deep learning models that have been around for a while. Its name comes from the fact that it is made up of numerous deep independent networks, each with its own collection of hidden units. Training, according to the DSN, is a collection of distinct training issues rather than a single issue. A mix of network and architectural aspects makes up the DSN. There are three modules in the DSN. Input, hidden layer, and output regions are all present in every component of the system. Using genuine input vectors and the prior layer's results, each module receives input from the preceding layer. To improve a module's efficiency and capability, DSN teaches it to work in isolation. For each unit, back-propagation is used, rather than the entire infrastructure, to train it. Conventional DBNs are a prevalent and practical network design because they outperform DSNs. [1]

LONG SHORT-TERM MEMORY

There are several uses for the LSTM (Long short-term memory) created by Hochreiter and Schmidhuber. For speech recognition, IBM chose LSTMs. Cells are memory components that may maintain their results for lengthy periods and are used by the LSTM because of their source. This makes it easy for the gadget to keep track of the most recent value. There are three gates in a memory cell that manage the flow of data in and out of the memory cell. [23,24]

- The gate, also known as the input terminal, regulates the flow of new data into the storage.
- The forgets port regulates gate is used to aid the cell in remembering new data when a previous bit of information is lost.
- The output gate's job is to regulate the records in the cell and use them as its result.

The weight of a cell can be utilized as a determining factor in a project. BPTT (Backpropagation through Time) is a weight-raising training strategy. Refinement is achieved by the use of network outages. Upgrade and reset gates are included in the Gated

Recurrent Unit (GRU). This cell's contents must be preserved, and an upgraded gate expresses that need. Cell contents are transferred to a new source using a reset gate. When the reset and upgrade gates are set to 1 and 0, a standard RNN is emulated by the GRU. The GRU model's capacity is lower than the LSTM models. As a result, it's expected to be more effective in terms of both implementation and learning. [25,26,27]

LSTM Gates Equations

$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i)$$

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f)$$

$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o)$$

Cell, candidate state, and output are:

$$\tilde{c}_t = tanh(w_c[h_{t-1}, x_t] + b_c)$$

$$C_t = f_t * c_{t-1} + i_t * \tilde{c}_t$$

$$h_t = o_t * tanh(c^t)$$

AUTO-ENCODER

There is a type of neural network known as an AE (Auto Encoder) that employs the backpropagation approach and learns without any supervision. The network places the initial input results to be the same as the outcome results. The network is trying to figure out a way to approximate a person's identification. It contains three layers: an input, an encoding layer (hidden unit), and an output layer (decoding unit). As the network seeks to recreate its input, the hidden unit is compelled to get the best descriptions of the input. The encoding layer specifies software that describes the input. Although not neural networks, auto-encoders are related to PCA (Principal Component Analysis). [28,29]When working with high-scalable data, auto-encoders are used, and data reduction specifies how a collection of data is expressed. The two primary structures utilized by Auto-encoder are Denoising Auto-encoder and Sparse Auto-encoder. In the Denoising Auto-encoder, they exploited noise dataset to assess the system weight, and in the Sparse Auto-encoder, they limited the activation state of decoding layers. An auto-encoder analyses the input before transferring it to an intrinsic change through nonlinear tracing. [30]

GENERATIVE ADVERSARIAL NETWORK

In 2014, Ian Goodfellow and colleagues at the University of Montreal initially developed Adversarial Networks (GAN). No matter what the field may be, GANs may imitate any type of data distribution. Unsupervised learning is used to train two models at once in GANs, which are a type of system. Variable numbers that are much lower than typical for the quantity of input that goes into training GANs are an important part of the GAN design process. Networks have to represent training data correctly, which improves their ability to create data that closely matches the training data. A discriminator (D) and a generator (G) comprise a GAN system. Discriminator encounters are used by the generator to learn how to produce a phony output that seems to be genuine [31]. As a discriminator, you must be able to detect the difference between what is a fake and what is a genuine piece of data.

$$\frac{min}{G}\frac{max}{D}V(D,G)$$

$$V(D,G) = E_{x\sim p_{data}(x)}[logD(x)] + E_{x\sim p_z(z)}[log(1 - D(G(z))]$$

GANs have three stages. To get started, the generator creates a random picture using numbers. As part of the real dataset, the discriminator gets this image as well as a stream of images. A second step is to use a discriminator to separate between real and fraudulent pictures, and then use that information to calculate a likelihood score between 0 and 1. For the generator to improve and train, the discriminator network gives feedback. When it comes to picture resolution, GANs have been utilized before, and they have the potential to be applied in several ways[32]. GAN's ability to create images from detailed captions, such as "a yellow truck with white doors" [33], is another key feature.

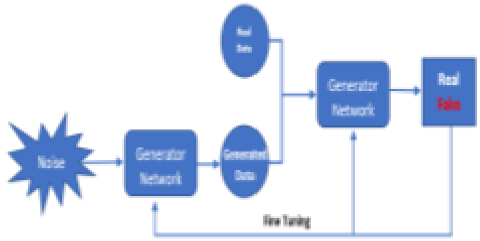

Figure5: *GAN Architectural Framework Pipeline*

APPLICATIONS OF DEEP LEARNING

DL aims to solve the most complicated elements of the input layer by using many layers of presentation. This unique technique to ML has already helped facial, voice, image, and text recognition systems, NLP, biomedical, and many other applications. Gradient descent and evolutionary algorithms have been used in recent research to reveal the model's optimization and fine-tuning. Deep learning technology is confronted with several challenges, including the need to scale computations, improve the parameters of deep neural networks, and develop and implement new approaches. This potential area of study is further hampered by the need to thoroughly examine a large number of sophisticated DNN (deep neural network) models. [34,35,36]

ADAPTIVE LEARNING

Chandra and Sharma [37] suggested including adjustable learning speeds and weight updates via the Laplacian score. Gaining weight and learning more quickly are attributed to the neurons. For example, error gradients-based techniques are conceivable for dealing with this. There are several ways to enhance the difficulty of finding a neuron's Laplacian score and incoming weights. The linear activation feature and the maximum out were used to achieve this on the benchmark datasets. Classification accuracy was improved as a result of the study. The Laplacian score could not be utilized online as a result of this method. They recommended utilizing the 'Exponential LP' with the Rectified Linear activation feature when data was available in streams and batches.

Xiao et al. [38] have developed a novel adaptive assessment method based on deep learning. Without the need for human interaction, these algorithms can extract characteristics from data. They used DNN, which has a greater percentage of success and failure prediction accuracy. They created two apps, partial testing and dynamic test ordering, both based on DNN capabilities. In addition to dynamically ordering tests, these programmers were used to making choices such as "pass or fail." The results of the tests showed an increase in precision and effectiveness.

AUTOMOTIVE INDUSTRY

Luckow et al.[39] offered automotive industry applications and methodology for deep neural networks. They concentrated their efforts on computer vision and CNNs. They used tagged datasets for their research. As a result of the paper's key contribution, consumers may now easily detect and evaluate automotive attributes. When they looked at both the training procedure and the classifiers' performance, they found that the taught classifiers performed well.

BIG DATA

Given today's fast growth in data size, the application has the potential to disrupt a wide number of sectors. With this, the path is cleared for big data forecasting and analytic solutions to meet previously unimaginable data and information needs. Deep learning, according to Chen and Lin [40], has serious drawbacks. Transformation solutions are required because of the huge size, diversity, chaotic labeling, and non-static distribution of the barriers. Many changes and potential for deep learning arose as a result of the current state of affairs surrounding big data. According to Gheisari et al. [41], a deep learning approach in big data analysis can uncover previously unknown and beneficial patterns at a high degree of abstraction that could not have been grasped before.

DEEP VISION SYSTEMS

Puthussery et al. developed a new machine learning-based technique for driver assistance systems [42]. As well as building a robot computer system and a system that leads itself to the marker depending on the object's position, they utilized CNN to identify the marker in the images. By utilizing a depth camera, they were able to gauge the distance between these landmarks and their location on the map. Using deep learning models created by Abbas et al. [43], real-time video processing applications may now be constructed. This is because deep learning algorithms are so strong in today's technological environment and may be used for live video processing applications, such as identifying, tracking, and recognizing video objects.

Their design and new method addressed processing costs, layer count, precision, and accuracy. They were successful. To effectively manage enormous amounts of video data using deep learning techniques, they built a strong, powerful framework with many elements called neurons. Unsupervised deep learning was used by Sanakoyeu et al. [44] to develop a system for identifying visual characteristics. A novel problem was proposed for detecting groups of samples that have a common ancestor, based on inadequate local similarity estimates. Disappointing connections were dispersed throughout several batches, all of which included identical materials. To offer a single representation for all of the data without tags, a CNN was utilized to make associations inside and across groups. In terms of posture analysis and object categorization, the proposed technique is superior to existing methods.

HEALTH CARE

Loh and Then [45] described a method for diagnosing and treating heart disease in rural areas, as well as the advantages, drawbacks, and solutions that come with implementing deep learning techniques. The need for medical software and telehealth services was particularly acute in rural regions. Portable medical gadgets and mobile computing, among other things, have been developed to address the problems that arise often in distant locations. Medical pictures can also benefit from the use of computer-aided design approaches. Physicians and patients will both gain from the use of the machine and deep learning techniques. Healthcare services in neglected regions may be accelerated by mobile technology advancements.

By utilizing artificial intelligence, Dai and Wang[46] propose a method for reducing the overburdening of doctors and nurses in the healthcare industry. Because of this, they concluded that using deep neural networks to detect health status was possible using pattern matching approaches and the depths recognition module. An inference graph-based use process module and a simulation environment with a body simulator to assist the body in treatment and a health-state simulator to alter the patient's health condition were also created by this group of researchers as a consequence, a body simulation module, a physical feature detection module, and a reform module utilizing Bayesian inference graphs were all constructed by the team. With additional statistical data, the research was shown to be the most successful. There were nine body basic kinds to choose from in the dataset utilized for the experiment.

TEXT SUMMARIZATION

Zhong et al.[47] developed a new query-oriented method for multi-document summarization based on deep learning methods. Dynamic programming was employed to keep track of and disclose the extraction capabilities in actual use. Concept extraction, summary creation, and reconstruction validity all fall under the purview of the model. It was therefore necessary to alter the deep structure to reduce data loss during reconstruction validation. For industrial uses, they were the right structure and did not require any training. The DUC 2005, 2006, and 2007 datasets were all used in the analysis. They were able to demonstrate that this method outperformed other methods of extraction. At each phase of the training process, Azar and Hamey [48] used an unsupervised auto-encoder to extract features that did not include the query at all, had a restricted local vocabulary, and lowered training and deployment costs. Studying emails from SKE and BC3, they found that AE had a more distinct function and increased basic word frequency.

DATA FLOW GRAPHS

To better comprehend data flow graph-based machine learning architectures, Abadi et al. [49] proposed the TensorFlow Graph, a component of the TensorFlow machine intelligence framework. The interactive diagram's conventional layout was constructed using serial updates. They used the original code's stepwise structure and the coupling of non-critical nodes to build the clustered network. To improve the growth's responsiveness and dependability, they also worked on the modular design and edge packing. Ripoll et al. [50] suggested a novel deep neural network-based filtering approach. The procedure was created to determine if a patient in the emergency room should be sent to cardiology for further evaluation. Using raw ECG values from 1320 people, the approach was tested. Algorithms included k-nearest neighbor, categorization, and learning machines. Support vector machines with Gaussian kernels were utilized in the experiment, and the results showed an accuracy of 84%, a sensitivity of 94%, and a specificity of 73%. It was proposed by Looks et al. [51] to merge multiple different input graphs with different shapes and nodes into a single input network known as dynamic batching. Dynamic computing graphs of varying sizes and forms were made more easily accessible via the development of new technological tools. To construct batch forms for various models, the group has created a library.

PLANT CLASSIFICATION

According to Lee et al. [52] deep learning techniques may be used to categorize plants. It aids botanists in making accurate and speedy species identifications. They used CNN and Deep Networks to extract significant leaf attributes from raw leaf data representations. To increase the plant classification system's capacity to discern species apart, researchers conducted extensive studies and developed hybrid feature and extraction models.

FAULT DETECTION IN POWER SYSTEM

By Wang et al., a new strategy for identifying power network faults was presented [53]. Finding and processing relevant data from an enormous amount of unlabeled data was the primary goal. The technique addressed data availability, local optimal improvement, and gradient dispersion. SCADA data from the power supply administration was used to monitor the power system. A layered auto-encoder (SAE) was used to train using missing features detected in several dimensions after the data was collected, preprocessed, and submitted. Afterward, a trained SAE is used for initialization, with the identifier verifying the kind of therapy. The outcomes confirmed the procedure's accuracy and viability. Based on CNN, Rudin et al. [54] presented one way of identifying power system breakdowns. As part of their investigation, they sought to categorize defective power system voltage signal samples. A two-way system for the dataset is built and three steps are taken to evenly distribute the workload. The voltage signal is measured at the line's commencement and the intersection of the two buses. The issue happened in the center of the line, in between two buses. Wavelet transform is used for feature extraction.

SOCIAL APPLICATION

Sentiment analysis frequently makes use of deep learning algorithms. Araque et al. [55] explained surface-based deep learning, which was based on manually generated features. They employed a deep learning-based sentiment classifier and a linear machine learning method for this experiment. With the classifier, it is possible to compare the results. There were two models for integrating the baseline and surface classifiers that they created. Microblogging was used to collect seven public datasets. The models' performances were highlighted in this video.

CONCLUSION

Deep learning is one of the most effective uses of machine learning. Deep learning algorithms' success and adaptability have been demonstrated in a variety of disciplines through their quick implementation. With its impressive accomplishments in deep learning and high accuracy rates, the technology demonstrates its usefulness while showing its progress and potential for further study and improvement. Making a superb deep learning app necessitates careful attention to the layer structure and learning oversight. A hierarchy is necessary for effective data classification and the supervisor is aware of how crucial it is to keep the database in excellent condition.

Improved machine learning applications, as well as a novel hierarchical layer processing technique, form the basis of deep learning. Digital picture and speech recognition are only two of the many uses for deep learning. Deep learning may be utilized as a preventive strategy in the present and the future because of the integration of face detection and voice recognition. Additionally, digital imaging is a subject that may be put to use in a variety of different areas of study. As a result of its success, deep learning is a hot issue in artificial intelligence right now. Finally, we can say that the use of deep learning in almost every application is gaining pace as data and computer power become more readily available. I believe that deep learning will achieve its aims and achieve new heights of success and happiness in a variety of applications over the next several years, including natural language processing (NLP) and object tracking (OT).

REFERENCES

1. Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2019). A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. Archives of Computational Methods in Engineering. doi:10.1007/s11831-019-09344-w
2. Soniya, Paul S, Singh L (2015) A review on advances in deep learning. In: Proceedings of IEEE workshop on computational intelligence: theories, applications and future directions (WCI), pp 1–6. https://doi.org/10.1109/wci.2015.7495514
3. Hewamalage H, Bergmeier C, Bandara K. Recurrent neural networks for time series forecasting: current status and future directions. Int J Forecast. 2020;37(1):388–427.
4. Jiang Y, Kim H, Asnani H, Kannan S, Oh S, Viswanath P. Learn codes: inventing low-latency codes via recurrent neural networks. IEEE J Sel Areas Inf Theory. 2020;1(1):207–16.
5. John RA, Acharya J, Zhu C, Surendran A, Bose SK, Chaturvedi A, Tiwari N, Gao Y, He Y, Zhang KK, et al. Optogenetics inspired transition metal dichalcogenide memristors for in-memory deep recurrent neural networks. Nat Commun. 2020;11(1):1–9.
6. Batur Dinler Ö, Aydin N. An optimal feature parameter set based on gated recurrent unit recurrent neural networks for

speech segment detection. Appl Sci. 2020;10(4):1273.

7. Jagannatha AN, Yu H. Structured prediction models for RNN based sequence labeling in clinical text. In: Proceedings of the conference on empirical methods in natural language processing. conference on empirical methods in natural language processing, vol. 2016, NIH Public Access; 2016. p. 856.
8. Pascanu R, Gulcehre C, Cho K, Bengio Y. How to construct deep recurrent neural networks. In: Proceedings of the second international conference on learning representations (ICLR 2014); 2014.
9. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics; 2010. p. 249–56.
10. Gao C, Yan J, Zhou S, Varshney PK, Liu H. Long short-term memory-based deep recurrent neural networks for target tracking. Inf Sci. 2019;502:279–96.
11. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84–90.
12. Zhou DX. Theory of deep convolutional neural networks: downsampling. Neural Netw. 2020;124:319–27.
13. Jhong SY, Tseng PY, Siriphockpirom N, Hsia CH, Huang MS, Hua KL, Chen YY. An automated biometric identification system using CNN-based palm vein recognition. In: 2020 international conference on advanced robotics and intelligent systems (ARIS). IEEE; 2020. p. 1–6.
14. Al-Azzawi A, Ouadou A, Max H, Duan Y, Tanner JJ, Cheng J. Deepcryopicker: fully automated deep neural network for single protein particle picking in cryo-EM. BMC Bioinform. 2020;21(1):1–38.
15. Wang T, Lu C, Yang M, Hong F, Liu C. A hybrid method for heartbeat classification via convolutional neural networks, multilayer perceptrons, and focal loss. PeerJ Comput Sci. 2020;6:324.
16. Li G, Zhang M, Li J, Lv F, Tong G. Efficient densely connected convolutional neural networks. Pattern Recogn. 2021;109:107610.
17. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J, et al. Recent advances in convolutional neural networks. Pattern Recogn. 2018;77:354–77.
18. Fang W, Love PE, Luo H, Ding L. Computer vision for behavior-based safety in

construction: a review and future directions. Adv Eng Inform. 2020;43:100980.
19. Palaz D, Magimai-Doss M, Collobert R. End-to-end acoustic modeling using convolutional neural networks for hmm-based automatic speech recognition. Speech Commun. 2019;108:15–32.
20. Li HC, Deng ZY, Chiang HH. Lightweight and resource-constrained learning network for face recognition with performance optimization. Sensors. 2020;20(21):6114.
21. Hubel DH, Wiesel TN. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. J Physiol. 1962;160(1):106.
22. Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning, vol. 1. Cambridge: MIT Press; 2016.
23. Boden, M.: A guide to recurrent neural networks and backpropagation. The Dallas project (2002)
24. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. 9(8), 1735–1780 (1997)
25. Liang, N.-Y., Huang, G.-B., Saratchandran, P., Sundararajan, N.: A fast and accurate online sequential learning algorithm for feedforward networks. IEEE Trans. Neural Netw. 17(6), 1411–1423 (2006)
26. Hosseini, M.-P., Hajisami, A., Pompili, D.: Real-time epileptic seizure detection from EEG signals via random subspace ensemble learning. In: 2016 IEEE International Conference on Autonomic Computing (ICAC), pp. 209–218. IEEE (2016)
27. Hosseini, M.P., Lau, A., Lu, S., Phoa, A.: Deep learning in medical imaging, a review. IEEE Rev. Biomed. Eng. (2019)
28. Hosseini, M.-P., Soltanian-Zadeh, H., Elisevich, K., Pompili, D.: Cloud-based deep learning of big EEG data for epileptic seizure prediction. In: 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp. 1151–1155. IEEE (2016)
29. Liao, D., Lu, H.: Classify autism and control based on deep learning and community structure on resting-state fMRI. In: 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), pp. 289–294. IEEE (2018)
30. Hosseini, M.-P.: A cloud-based brain-computer interface to analyze medical big data for epileptic seizure detection. In: The 3rd Annual New Jersey Big Data Alliance (NJBDA) Symposium (2016)
31. Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A.: Generative adversarial networks: an overview.

IEEE Signal Process. Mag. 35(1), 53–65 (2018)
32. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4681–4690 (2017)
33. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., Lee, H.: Generative adversarial text to image synthesis (2016). arXiv preprint arXiv:1605.05396
34. Chea P, Mandell JC. Current applications and future directions of deep learning in musculoskeletal radiology. Skeletal Radiol. 2020;49(2):1–15.
35. Kuutti S, Bowden R, Jin Y, Barber P, Fallah S. A survey of deep learning applications to autonomous vehicle control. IEEE Trans Intell Transp Syst. 2020;22:712–33.
36. Jiao L, Zhang F, Liu F, Yang S, Li L, Feng Z, Qu R. A survey of deep learning-based object detection. IEEE Access. 2019;7:128837–68.
37. Chandra B, Sharma RK (2016) Deep learning with adaptive learning rate using Laplacian score, expert systems with applications. Int J 63(C):1–7
38. Xiao B, Xiong J, Shi Y (2016) Novel applications of deep learning hidden features for adaptive testing. In: Proceedings of the 21st Asia and South Pacific design automation conference, pp 743–748
39. Luckow A, Cook M, Ashcraft N, Weill E, Djerekarov E, Vorster B (2017) Deep learning in the automotive industry: applications and tools. In: Proceedings of the IEEE international conference on big data, pp 3759–3768
40. Chen XW, Lin X (2014) Big data deep learning: challenges and perspectives. IEEE 2:514–525. https://doi.org/10.1109/ACCESS.2014.2325029
41. Gheisari M, Wang G, Bhuiyan MZA (2017) A survey on deep learning in big data. In: Proceedings of the IEEE international conference on embedded and ubiquitous computing (EUC), pp 1–8
42. Puthussery AR, Haradi KP, Erol BA, Benavidez P, Rad P, Jamshidi M (2017) A deep vision landmark framework for robot navigation. In: Proceedings of the system of systems engineering conference, pp 1–6
43. Abbas Q, Ibrahim MEA, Jafar MA (2018) A comprehensive review of recent advances on deep vision systems. Artif Intell Rev. https://doi.org/10.1007/s10462-018-9633-3

44. Sanakoyeu A, Bautista MA, Ommer B (2018) Deep unsupervised learning of visual similarities. Pattern Recogn 78:331–343
45. Loh BCS, Then PHH (2017) Deep learning for cardiac computer-aided diagnosis: benefits, issues & solutions. M Health. HTTPS://doi.org/10.21037/mhealth.2017.09.01
46. Dai Y, Wang G (2018) A deep inference learning framework for healthcare. Pattern Recogn Lett. https://doi.org/10.1016/j.patrec.2018.02.009
47. Zhong SH, Li Y, Le B (2015) Query oriented unsupervised multi-document summarization via deep learning. Expert Syst Appl, pp 1–10
48. ar MY, Hamey L (2017) Text summarization using unsupervised deep learning. Expert Syst Appl 68:93–105
49. Abadi M, Paul B, Jianmin C, Zhifeng C, Andy D, Jefrey D, Matthieu D (2016) TensorFlow: a system for large-scale machine learning. In: The proceedings of the 12th USENIX symposium on operating systems design and implementation (OSDI'16), vol 16, pp 265–283
50. Ripoll VJR, Wojdel A, Romero A, Ramos P, Brugada J (2016) ECG assessment based on neural networks with pertaining. Appl Soft Comput 49:399–406
51. Looks M, Herreshof M, Hutchins D, Norvig P (2017) Deep learning with dynamic computation graphs. In: Proceedings of the international conference on learning representation, pp 1–12
52. Lee SH, Chan CS, Mayo SJ, Remagnino P (2017) How deep learning extracts and learns leaf features for plant classification. Pattern Recogn 71:1–13
53. Wang Y, Liu M, Bao Z (2016) Deep learning neural network for power system fault diagnosis. In: Proceedings of the 35th Chinese control conference, 1–6
54. Rudin F, Li GJ, Wang K (2017) An algorithm for power system fault analysis based on convolutional deep learning neural networks. Int J Res Educ Sci Methods 5(9):11–1
55. Araque O, Corcuera-Platas I, Sánchez-Rada JF, Iglesias CA (2017) Enhancing deep learning sentiment analysis with ensemble techniques in social applications. Expert Syst Appl 77:236–246