



# An Extensive Study on Deep Learning: Techniques, Applications

Ruchi Mittal<sup>1</sup> · Shefali Arora<sup>2</sup> · Varsha Bansal<sup>1</sup> · M. P. S. Bhatia<sup>2</sup>

Received: 24 May 2020 / Accepted: 10 January 2021 / Published online: 3 February 2021  
© CIMNE, Barcelona, Spain 2021

## Abstract

Deep learning is associate degree future field of machine learning (ML) analysis. It consists of variety of numerous concealed layers of artificial neural networks (ANN). Deep learning methods apply nonlinear transformation and high-level model abstraction to giant databases. Recent advances in deep learning design among several fields have already contributed considerably to computer science. This text presents a study on commitments and novel uses of escalated instruction of intensive education. The subsequent examination presents however and during which key application intensive algorithms for learning are used. Additionally, deep learning methodology is given with enhancements and its hierarchy in linear and non-linear functions and compared with a lot of ancient algorithms in widespread applications. The status of the survey of art provides a common rundown on the novel thought and therefore rising learning and deep learning quality.

## 1 Introduction

Artificial intelligence (AI), as an incontestable by machines, has been known to be an efficient approach to human learning and reasoning. In 1950, "The Alan Mathison Turing Test" was planned for lovely rationalization of how a pc will perform human psychological feature reasoning. As an exploration space, AI has a lot of functional analysis subfields. For instance: the linguistic communication process (NLP) will enhance the writing expertise in terms of grammatical structure and writing system mistakes [1, 2]. Therefore, classic subdivision among informatics is AI that is known as translation between languages. Recently, machine learning and data processing have become the focus of attention and square measure the foremost common topic among the analysis community. These combined fields of study evaluate several prospects of the characterization of databases [3]. Through the years, with applied mathematics objectives, many databases are collected. Applied mathematics curves will describe past and predict future behavior. But throughout the last decades, solely classic techniques and algorithms have been used to method this knowledge, whereas improvement will lead to an efficient self-learning.

A higher decision is the application of supported existing values, the use of many criteria, and advanced strategies of statistics. The first vital application of this improvement is in the medical field, where symptoms, causes, and medical solutions generate giant databases that are to predict higher treatment. Insight, as we probably are aware, is the capacity to get and apply information [4]. Information is the data gained through understanding. Experience is the information that increased through exposure (training). Summarizing the terms, we get human-made brainpower as the "duplicate of something natural (i.e., people) 'WHO' is equipped for obtaining and applying the data it has increased through the presentation." AI uses many tools, which include renditions of search and numerical streamlining, rationale, techniques dependent on likelihood, and financial matters. The AI field draws upon software engineering, arithmetic, brain research, semantics, reasoning, neuroscience, fake brain science, and numerous others. Utilizations of AI incorporate Natural Language Processing, Gaming, Speech Recognition, Vision Systems, Healthcare, Automotive, and so on [5].

An AI framework is made from an operator and its condition. An agent (e.g., human or robot) is whatever can see its condition through sensors and follows up on that condition through effectors [6]. Canny specialists must have the option to set objectives and accomplish them. In traditional arranging issues, the specialist can accept that it is the initial framework acting on the planet, permitting the operator to be sure of the outcomes of its activities. In any case, if the specialist isn't the main on-screen character, at that point, it

✉ Ruchi Mittal  
ruchi.mittal138@gmail.com

<sup>1</sup> Ganga Institute of Technology and Management, Jhajjar, Haryana, India

<sup>2</sup> Netaji Subhas Institute of Technology, New Delhi, India

necessitates that the operator can reason under vulnerability. Which requires an operator that would not only survey its condition and make forecasts yet, besides, assess its expectations and adjust dependent on its evaluation. Normal language handling enables machines to peruse and comprehend human language. Some apparent uses of common language preparation incorporate data recovery, content mining, question noting, and machine interpretation. Machine discernment is the capacity to utilize contribution from sensors (for example, cameras, mouthpieces, sensors, and so on.) to find parts of the world. e.g., Computer Vision. Ideas, for example, game hypothesis, choice hypothesis, require that an operator have the option to recognize and demonstrate human feelings.

Typically, understudies get dumbfounded between Machine Learning and Artificial Intelligence. Machine learning, a significant idea of AI investigation since the field's start, is the assessment of PC tallies that often improve through appreciation. The consistent evaluation of AI checks and their show is a bit of theoretical programming structuring known as a computational learning hypothesis.

Stuart Shapiro detaches AI assessment concerning three methodologies, which he calls: computational brain requests about, computational viewpoint, and programming designing [7]. Computational psyche science is utilized to make PC programs that impersonate human straightforwardly. Computational point of view is utilized to build up an adaptable, free-spilling PC mind. Executing programming structuring serves the objective of making PCs that can perform undertakings that no one but individuals could beforehand achieve.

Conspicuous cases of AI fuse independent vehicles, (for instance, robots and self-driving automobiles), clinical assurance, making craftsmanship, (for instance, refrain), exhibiting logical speculations, messing around, (for instance, Chess or Go), web lists, (for instance, Google search), online associates, (for instance, Siri), picture affirmation in photographs, spam isolating, conjecture of lawful choices and concentrating on online takes note. Various applications join Healthcare, Automotive, Money, Video games thus on [8–11]. Are there cutoff points to how keen machines—or human–machine crossbreeds—can be? A genius, hyperintelligent, or superhuman insight is a theoretical operator that would know far outperforming that of the most brilliant and most talented human psyche. "Genius" may likewise allude to the structure or level of knowledge controlled by such a specialist.

The term Machine Learning was composed by Arthur Samuel in 1959, an American pioneer in the field of PC gaming and human-made intellectual prowess, and communicated that "it empowers PCs to learn without being unequivocally changed."

Furthermore, in 1997, Tom Mitchell gave an "inside and out introduced" logical and standard definition that "A PC program is said to pick up sincerely E concerning some endeavor T and some show measure P, if its introduction on T, as evaluated by P, improves with experience E.

## 1.1 Classification of Machine Learning

Machine learning is arranged into three significant classes, contingent upon the idea of the learning "sign" or "reaction" accessible to a learning framework which is as per the following:-

1. *Supervised learning*: At the point when estimation gains from model information and related objective reactions that can include numeric attributes of string names, for example, classes or stamps, to later predict the right answer when given new models go under the portrayal of Supervised learning. This methodology is to be fulfilled like AI under the supervision of an instructor. The educator provides formal guides to the understudy to hold, and the understudy by then gets general guidelines from these models [12].
2. *Unsupervised learning*: This kind of computation will when all is said in done revamp the data into something other than what's expected, for instance, new features that may address a class or another course of action of un-related characteristics. They are entirely important in giving individuals encounters into the hugeness of data and new accommodating commitments to oversee AI estimations [13].  
As a kind of learning, it takes after the technique's individuals use to comprehend that things or events are from a comparative class, for instance, by watching the degree of likeness between objects. Some proposal structures that you find on the web through publicizing robotization rely upon this sort of learning.
3. *Reinforcement learning*: Right when you present the computation with models that need names, as in independent learning. In any case, you can go with a model with a positive or negative contribution by the course of action the figuring proposes under the class of Reinforcement acknowledging, which is related with applications for which the computation must choose (so the thing is prescriptive, not just captivating, as in solo learning). The decisions bear outcomes [14]. In the human world, it is a lot of equivalent to learning by experimentation.

Blunders assist you with learning since they have a punishment included (cost, loss of time, lament, torment, etc.), instructing you that a specific game-plan is less inclined to prevail than others. A fascinating case of

fortification learning happens when PCs figure out how to play computer games without anyone else.

Right now, application presents the calculation with instances of explicit circumstances, for example, having the gamer stuck in a labyrinth while staying away from an adversary. The application tells the calculation the result of moves it makes, and learning happens while attempting to keep away from what it finds to be risky and to seek after endurance. You can view how the organization Google DeepMind has made a support learning program that plays old Atari's computer games. When viewing the video, notice how the program is at first awkward and incompetent yet consistently improves with preparing until it turns into a hero [15–17].

4. *Semi-supervised learning*: Where an insufficient getting ready sign is given, it is a readiness set with a couple (routinely large quantities) of the target yields missing. There is a remarkable case of this standard known as Transduction, where the entire course of action of issue events is known at learning time. On the other hand, some segments of the destinations are absent [18].

## 1.2 Categorizing Based on Required Output

Another course of action of AI undertaking develops when one considers the perfect yield of a machine-learned framework [19]:

1. *Classification* When data sources are divided into at any rate two classes, and the understudy must make a model that delegates covered commitments to in any event one (multi-name request) of these classes. This is usually taken care of in a coordinated way. Spam isolating is an instance of collection, where the data sources are email (or other) messages, and the classes are "spam" and "not spam."
2. *Regression* This is furthermore a coordinated issue, a circumstance when the yields are steady rather than discrete.
3. *Clustering* When a great deal of data sources is to be parceled into social events. Not in the least like all together, the social occasions are not known up until now, making this customarily an independent task.

## 1.3 Data in Machine Learning

### 1.3.1 Information

It may be any ordinary reality, respect, substance, sound, or picture that isn't being deciphered and isolated. It is the most fundamental piece of all Data Analytics, Machine Learning, Artificial Intelligence. Without information, we can't set up any model, and all cutting edge research and

computerization will go vain. Large Enterprises are encountering heaps of cash to gather a great deal of explicit information, at any rate, as could be ordinary.

### 1.3.2 Example

For what reason did Facebook secure WhatsApp by completing at a huge cost of \$19 billion?

The best possible response is staggeringly essential and educated – it is to advance toward the clients' data that Facebook presumably won't have yet WhatsApp will have. This data of their clients is of critical centrality to Facebook as it will engage the undertaking of progress in their associations.

### 1.3.3 Data

It is the information that has been deciphered and controlled. It also now has some noteworthy induction for the customers.

### 1.3.4 Information

Information that has been deciphered and controlled and now has some noteworthy deduction for the customers.

### 1.3.5 Data Preparation

The bit of information we use to set up our model. This is the information that your model genuinely sees (both data and yield) and increments from.

### 1.3.6 Approval Data

The piece of information, which is utilized to do a continuous assessment of model fit on preparing dataset and improving hyperparameters. This information has its influence when the model is making.

### 1.3.7 Testing Data

For the entirely orchestrated model, testing information gives the unbiased assessment. Right when we feed in the duties of Testing information, our model will predict some values (without seeing real yield). After want, we assess our model by separating it, and confirmed return present in the testing information, and these are the strategies by which we survey and perceive how much our model has gotten from the encounters feed in as arranging information, set at the hour of preparing.

## 1.4 Properties of Data

- *Volume* Size of Data. With developing total populace and innovation at introduction, colossal information is being created every single millisecond [20].
- *Variety* Various types of information—social insurance, pictures, recordings, sound clippings.
- *Velocity* Pace of information spilling and age.
- *Value* Importance of information as far as data which scientists can construe from it.
- *Veracity* Assurance and accuracy in information we are dealing with.

## 2 Fundamental Concepts

Deep learning calculations square measure underpins appropriated portrayals. The fundamental supposition behind scattered descriptions is the cooperation of things creates that found information in layers. Deep learning misuses this thought of important hierarchic factors any place more significant level, many conceptual ideas gained from lower stages [21]. These structures square measure shapes with an insatiable layer-by-layer system. Profound learning enables that alternatives to square gauge accommodating for learning. Inside the instance of administered learning undertakings, escalated learning techniques make an interpretation of information into center designing, convert reduced middle of the road portrayals into main parts, and convey the great had relations with structures that conquer repetition in representations. A few profound learning calculations are applied to new learning errands. This can be an indispensable preferred position because of unlisted information square measure regularly a great deal of broad than marked information. Partner degree case of a profound structure which will be prepared to utilize such techniques could be a profound conviction arrange. Deep neural systems are typically deciphered: all-inclusive estimation hypotheses or chance approximations [22]. The widespread guess hypothesis gives the adaptability to take care of feed-forward neural networks with a shrouded layer of constrained structure for elements of uniform size. The essential verification for sigmoid enactment capacities was uncovered by Siegenko in 1989 and summed up to the feed-forward multi-layer structure in 1991 by Hornick [23] (Fig. 1).

### 2.1 Types of deep learning approaches

Deep learning can be divided into these categories: Supervised, semi-supervised or partially supervised, and unsupervised. Besides this, reinforcement learning is another category of deep learning, which could be unsupervised or semi-supervised (Fig. 2).

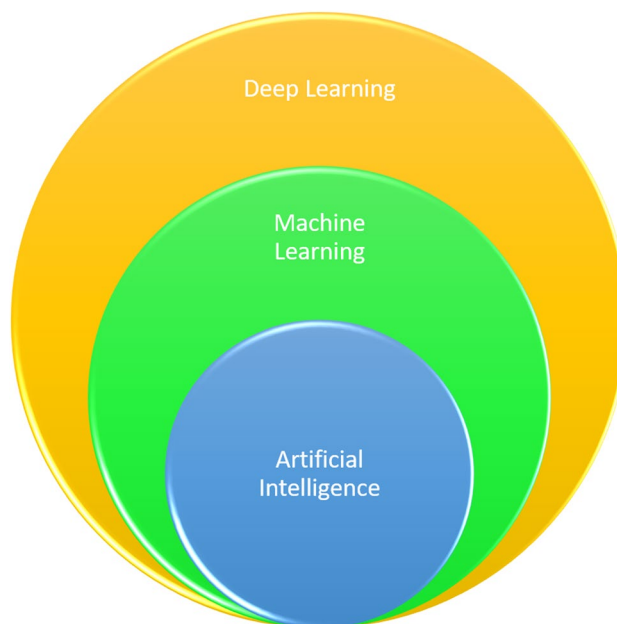


Fig. 1 General structure for deep learning

#### 2.1.1 Deep Supervised Learning

Supervised learning makes use of labeled data i.e. there is a set of inputs and corresponding outputs or labels. Based on the model, which is trained, we can predict values of unseen and new data. Model parameters can be modified in order to get better outputs. Different deep learning approaches for supervised learning are: Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), including Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU). The various studies which make use of these approaches are summarized in the table:

Study	Technique used	Application
Szedegy et al. [24]	Deep neural networks	Object detection
Kombrink et al. [25]	Deep neural networks	Language modeling in meeting recognition
Hinton et al. [26]	Deep neural networks	Automatic speech recognition
Chung et al. [27]	Recurrent neural networks	Multiscale RNN network
Sainath et al. [28]	Convolutional neural networks	Automatic speech recognition
Martin et al. [29]	Deep neural networks	Automatic story generation
Ren et al. [30]	Convolutional neural networks	Faster real time object detection

### 2.1.2 Deep Semi-supervised Learning

Semi supervised learning is based on datasets which are partially labeled. Popular such techniques are DRL and Generative Adversarial Networks (GAN) GAN is discussed in Sect. 7. Section 8 surveys DRL approaches. Additionally, RNN, including LSTM and GRU, are used for semi-supervised learning as well. Some of the researched done in this domain are summarized in the table:

Study	Technique used	Application
Miyato et al. [31]	Adversarial learning	Virtual adversarial training applicable to semi supervised learning
Wang et al. [32]	Classification	Autoencoding transformations for semi supervised learning
Gong et al. [33]	Classification	Semi supervised image classification

### 2.1.3 Deep Unsupervised Learning

Unsupervised learning works without the use of labels. The agent itself learns the important relationships and structures within the input data. Some of the popular techniques are Auto-Encoders (AE), Restricted Boltzmann Machines (RBM), and the recently developed GAN. In addition, RNNs, such as LSTM and RL, are also used for unsupervised learning in many application domains. Sections 6 and 7 discuss RNNs and LSTMs in detail.

Study	Technique used	Application
Tang et al. [34]	GANs	Multimodal image translation
Royer et al. [35]	Adversarial autoencoder	Mapping of source to target image
French et al. [36]	Ensemble methods	Gradient descent with exponential moving average
Saito et al. [37]	Asymmetric tri training	Prediction of true label based on confidence
Hung et al. [38]	generative adversarial learning	Semantic segmentation

### 2.1.4 Deep Reinforcement Learning

Deep Reinforcement Learning began in 2013 with Google Deep Mind [5, 6]. Since then, various studies have been

done in RL. Based on sample inputs, the agent predicts a value and gets a reward or penalty based on its move. Given  $P$  is an unknown probability distribution, the agent is asked a question by the environment and it gives a noisy answer. This approach could be semi-supervised as well. There are many semi-supervised and un-supervised techniques that have been implemented based on this concept (in Sect. 8).

Fortification learning could be a subfield of AI during which frameworks square measure prepared by accepting virtual "prizes" or "disciplines". Google's DeepMind made use of fortification by finding how to make human victors lose in Go games. Support learning is also used in computer. Some important calculations in this domain are:

- Q-learning.
- Deep Q arranges.
- State-Action-Reward-State-Action (SARSA).
- Deep Deterministic Policy Gradient (DDPG).

Fortification learning is a zone of supervised as well as unsupervised machine learning, as it involves the use of required moves to boost the reward which would be claimed for a specific move. This is done by finding the most ideal conduct applicable in a circumstance.

The primary concerns in Reinforcement learning are:

- *Input* The input should be obtained from the underlying state.
- *Output* There are numerous conceivable outputs as there are an assortment of answers for a particular action.
- *Training* This depends on information available, and the model will restore the state. The client will either remunerate or rebuff the model depending on the output.
- The model keeps on learning.
- The best arrangement is chosen dependent on the greatest prize.

There are two kinds of reinforcement learning:

**2.1.4.1 Positive** Uplifting feedback is characterized as when an occasion, happens because of specific conduct, expands the quality and the recurrence of the conduct. At the end of the day, it positively affects the conduct.

Favorable circumstances of support learning are:

- Maximizes Performance.
- Sustain Change for an extensive stretch of time.

Detriments of support learning:



- Too much Reinforcement can prompt over-burden of states which can decrease the outcomes.

**2.1.4.2 Negative** Negative Reinforcement is characterized as fortifying of conduct on the grounds that a negative condition is halted or dodged.

Focal points of support learning:

- Increases behavior.
- Provide disobedience to the least standard of execution.

Disservices of support learning:

- It only gives enough to get together the base conduct.

Different Practical uses of Reinforcement Learning.

- RL can be utilized in mechanical technology for modern robotization.
- RL can be utilized in AI and information handling.
- RL can be utilized to make preparing frameworks that give custom guidance and materials as per the necessity of understudies.

RL can be utilized in huge conditions in the accompanying circumstances:

1. A model of the earth is known, yet a scientific arrangement isn't accessible.
2. Only a reenactment model of the earth is given (the subject of recreation-based optimization) [6];
3. The best way to gather data about the earth is to communicate with it.

It is harder to learn as compared to supervised algorithms as there is no straight forward loss function. The main differences are that interaction is done in a state-based environment and queries are done through interactions. We do not have full access to the function we are trying to optimize.

Based on the type of problems and parameters involved, we can decide what type of RL algorithm needs to be used. DRL is the best technique when there are a lot of parameters. If the problem has fewer parameters for optimization, a derivation free If RL approach is good. An example of this is annealing, cross entropy methods, and SPSA.

Study	Technique used	Application
Hasselt et al. [39]	DRL	Double Q learning
Hausknecht and Stone [40]	DRL	Deep recurrent Q learning

Study	Technique used	Application
Hessel et al. [41]	DRL	Improvements in deep reinforcement learning
Wu et al. [42]	DRL	Trust region method
Dabney et al. [43]	Distributed DRL	Use of Quantile Regression

### 3 Deep Learning Process

Deep learning, otherwise called the Deep neural system, is one of the ways to deal with AI. Other significant methodologies incorporate choice tree learning, inductive rationale programming, grouping, fortification learning, and Bayesian networks. Deep learning is an extraordinary sort of AI. It includes the investigation of ANN and ML related calculations that contain more than one covered up layer. Deep learning includes scientific demonstrating, which can be thought of as a structure of straightforward squares of a specific sort, and where a portion of these squares can be changed in accordance with better anticipation of the last outcome. The word "profound" implies that the piece has a considerable lot of these squares stacked on one another—in a progression of expanding multifaceted nature. The yield gets produced by means of something many refer to as Backspread within a bigger procedure called Gradient drop which lets you change the parameters in a manner that improves your model [44].

Conventional AI calculations are straight. Deep learning calculations are stacked in a progressive system of expanding complexity. The capacity to process enormous quantities of highlights makes Deep learning ground-breaking when managing unstructured information. Be that as it may, Deep learning calculations can be pointless excess for less unpredictable issues since they expect access to an immense measure of information to be powerful. For example, ImageNet, the basic benchmark for preparing Deep learning models for far-reaching picture acknowledgment, approaches more than 14 million images. If the information is excessively straightforward or fragmented, it is extremely simple for a Deep learning model to become overfitted and neglect summing up of new information. Therefore, Deep learning models are not as successful as different methods, (for example, supported choice trees or direct models) for most reasonable business issues, for example, understanding client beat, distinguishing false exchanges and different cases with

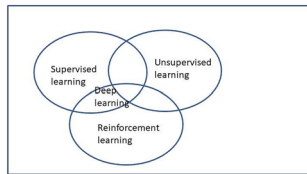


Fig. 2 Relationship between types of deep learning approaches

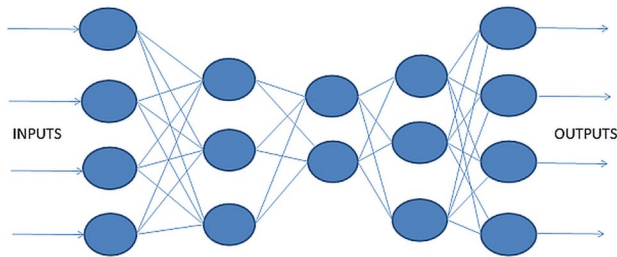


Fig. 3 Sample feedforward neural network

little datasets [45] and fewer highlights. In specific cases like multiclass arrangement, Deep learning can work for little, organized datasets.

A serious neural system gives dynamic exactness in a few undertakings, from object acknowledgment to discourse recognition. It will adapt precisely, unmistakably coded by software engineers while not predefined data. To get a handle on the idea of serious instruction, picture a family with a baby and a parent. The child focuses to things by the side of his finger and persistently says the word 'feline'. As her oldster's stress concerning her instruction, they keep on advising her 'Truly, she could be a feline' or 'No, she isn't a feline'. The child will in general reason to things, anyway, turns into a great deal of exact with 'feline'. The confined child where it counts, doesn't perceive why he will say whether it's a feline or not. She took in a manner to deliver propelled alternatives that escort a feline by watching pets comprehensively and still represent considerable authority in subtleties simply like the tail or nose before making up your brain. A neural system works likewise. Each layer speaks to a Deep degree of data, to be specific, the order of data. A neural system with four layers can get familiar with a great deal of cutting edge include than with two layers. Learning happens in two phases. The main stage comprises of applying a nonlinear change of the info and makes an applied arithmetic model as a yield. In the subsequent part targets rising the model with a scientific philosophy alluded to as result. The neural system rehashes these two stages a huge number of times till it arrives at a good degree of precision. This traditional dancing reiteration is named redundancy [46]. To give partner degree model, look at the proposition beneath, the model

making an endeavor to work out an approach to move. At the point when ten minutes of instructing, the model has no arrangement anyway the move is performed, and it's kind of a scrawl.

Arrangement of neural systems.

- *Shallow neural system* Single shrouded layer in the middle of information and yield.
- *Deep Neural Network* Deep neural system comprises more than one layer. For example, the Google LeNet model checks twenty-two layers for picture acknowledgment. These days Deep learning is utilized from multiple points of view like a driverless car, transportable, Google program, misrepresentation discovery, TV, at that point on.

## 4 Types of Deep Learning Networks

### 4.1 Feed-Forward Neural Networks

Feed-forward Neural Networks is the simplest form of artificial neural network. In these networks, inputs are fed to the input layer, followed by extraction of features using hidden layer. The final layer is output layer, which is used for classification or regression. Thus, output layer is the destination for the knowledge learned [47] (Fig. 3).

The main characteristics of these networks are:

1. Perceptron's have various layers, in which first layer takes the inputs and las layer gives the output. The middle layers are called hidden layers as they are used for feature extraction and have no connection with the outside world.
2. Every perceptron in one layer is connected to a perceptron in the next layer. As information is fed forward to the next layer, these networks are called feed forward networks.
3. There is no connection between perceptron's in the same layer.

Some of the studies related to feed forward networks are as follows:

Study	Application
SaishanmugaRaja and Rajagopalan [48]	Iris recognition
Tran et al. [49]	Analysis of data for financial cases
Qi et al. [50]	Neural estimators in aeronautic components
Huang et al. [51]	Detection of Insomnia from EEG and ECG

### 4.2 Recurrent Neural Networks (RNNs)

The Recurrent neural networks (RNN) is a neural network where output from the previous step is fed an input in the current step. Inputs and outputs are independent of each other in case of feedforward neural networks. However, in cases where we need to predict the next word, there is a need to include previous words in the input. Thus, RNN came into existence, which makes use of a hidden state to store such information. This network will store information in reference nodes, permitting it to be told knowledge sequences and output variety or the other sequence [52]. It is a synthetic neural network consisting of association loops between neurons.

RNN considers the input sequences to predict the next word in a sentence. These networks are called recurrent because this step is carried for every input. This helps us to predict the next word in the sentence. As these neural networks consider the previous word during predicting, it acts like a memory storage unit which stores it for a short period of time (Fig. 4).

RNN neurons can receive a symbol that points towards the start of a sentence. The network receives the word "do" as input, forming a vector of numbers. This vector is fed back to the nerve cell to supply memory to the network. At this stage, the network stores the word 'do', which was received in the first place. The network can likewise proceed to consecutive words. It takes the words "you" and "want". The position of neurons is updated on every occasion a word is received. The last step is when receiving

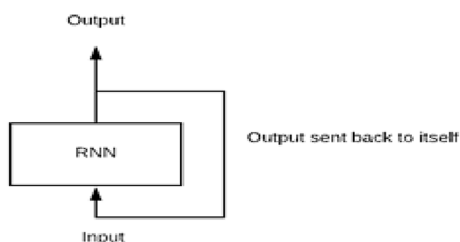
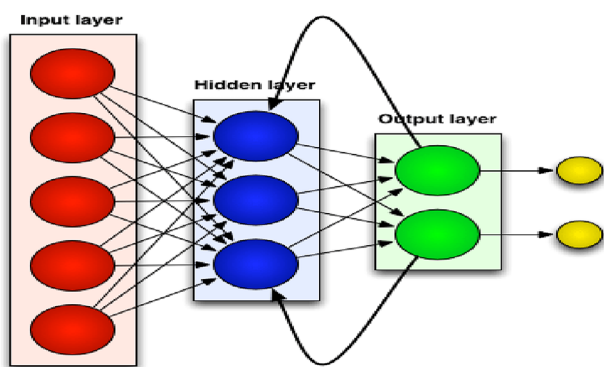


Fig. 4 Recurrent neural networks

the word "A". The neural network can give an opening for every English word which will be wont to complete a sentence. A well-trained RNN possible offers a high chance for "cafes," "drinks," "burgers," etc. [53].

Study	Application
Sydney Kasongo [54]	Wireless intrusion system
Bouktif [55]	Load forecasting
Ivan Zhang [56]	Predicting trend of dissolved oxygen
Farid Razzak [57]	Multimodal attention-based approach
Tong et al. [58]	Use of RNN and LSTM

### 4.3 Convolutional Neural Networks (CNN)

CNN could be a multi-layered neural network with a novel design designed to speedily extract advanced options of the information at every layer to search out the output. CNN square measure likeminded for sensory activity tasks [59], 60 (Fig. 5).

CNN is generally utilized once there's partner degree unstructured information set (e.g., pictures) and doctors found a workable pace from it. On the off chance that for example, if the assignment is to foresee the picture subtitle: CNN finds an image guess that a feline, this picture, in pc terms, could be a grouping of pixels. By and large one layer for the greyscale picture and three layers for a shading picture. During highlight learning (for example concealed layers), the system can set up alternatives, for example, feline's tail, ears, and so forth. At the point when the system has completely taken in a manner to recognize a picture, it will give an opening for each picture that it knows about. The name with the absolute best possibility can turn into the forecast of the system [61].

A Convolutional Neural Network (CNN) consists of at least one convolutional layer (frequently with a subsampling step), later followed by at least one completely associated layer in a standard multilayer neural system. Such networks exploit the 2D structure of an image. Features are extracted by finding association, followed by performing a

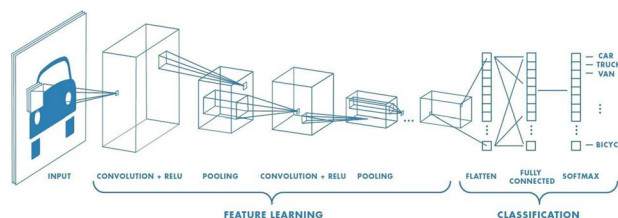


Fig. 5 Convolutional neural network



pooling operation. This brings about interpretation invariant highlights. The advantage of using CNNs is that they are simpler to prepare, and parameters are lesser as compared to completely associated systems with a similar number of shrouded units [62].

A CNN comprises of various convolutional and subsampling layers alternatively followed by completely associated layers. The contribution to a convolutional layer is a  $mm \times mm \times rr$  picture where  $mm$  is the tallness and width of the picture and  $rr$  is the number of channels, for example, an RGB picture has  $r=3$ . Given that there are  $kk$  filters in a convolutional layer, the final size will be  $mm-kk + 1 \times mm-kk + 1$ . This is followed by a mean or max pooling of size  $p \times pp \times p$  adjacent areas where  $p$  goes between 2 for little pictures (for example MNIST) and is generally not more than 5 for bigger information sources. A nonlinear activation function such as tanh or sigmoid is performed at each step over the component map. The figure beneath represents a full layer in a CNN comprising of convolutional and subsampling sublayers. Units of a similar shading have tied loads.

After the convolutional layers, there might be any number of completely associated layers. The thickly associated layers are indistinguishable from the layers in a standard multilayer neural system.

Study	Application
Kumar et al. [58]	Profession analysis using handwritten data
Chi et al. [63]	gender classification
Wang et al.[64]	Image forgery detection
Wang et al. [65]	Detection of multiple objects
Kong et al. [66]	Skin disease diagnosis using photographs

## 5 Applications of Deep Learning

Deep learning alludes to relate degree reflection layer investigation and hierarchic techniques. In any case, it is used in a few genuine applications [67]. As the partner degree model, among advanced picture handling, dark scale picture shading from a picture was done physically by clients United Nations organization needed to choose each shading bolstered their call. Shading is performed precisely by a pc by Actualizing a Deep learning algorithmic guideline. So also, stable is supplemental to quiet percussion recordings exploitation enduring Neural Networks (RNN) as a piece of severe learning systems. Deep learning is comprehended as an approach to help results and upgrade interim in a few registering forms. Inside the field of the etymological correspondence

process, escalated learning methodologies are applied for picture subtitle age and penmanship age. The ensuing applications square measure grouped into unadulterated advanced picture procedure, medicine and bioscience [68].

### 5.1 Automatic Discourse Acknowledgment

Enormous scope programmed discourse acknowledgment is that the first and premier thundering instance of Deep learning. The LSTM RNN will learn "Deep learning" assignments, including multi-second interims with discourse occasions isolated by many seconds of unmistakable time steps, any place a period step compares to around 10 ms. LSTM with Gates is serious with old discourse identifiers on certain undertakings [69]. SSS the data set contains 630 speakers from eight significant lingos of yank English, with each speaker perusing ten sentences. Its small size attempts a few designs. A ton of fundamentally, the TIMIT perform issues telephone succession acknowledgment, dislike word-arrangement acceptance, licenses for frail telephone composed word language models. This allows the quality of acoustic demonstrating parts of discourse acknowledgment to be broke down a ton of just. These underlying outcomes, just as the mistake rates recorded underneath and estimated as % telephone blunder rate (per), are outlined since 1991.

Deep learning applications are utilized in businesses from mechanized heading to clinical gadgets [70].

### 5.2 Automated Driving

Car scientists are utilizing Deep figuring out how to consequently distinguish articles, for example, stop signs and traffic lights. Likewise, in-depth learning is used to identify walkers, which helps decline mishaps [71].

### 5.3 Aerospace and Defense

Deep learning is utilized to distinguish objects from satellites that find regions of premium and recognize sheltered or hazardous zones for troops. Also, many departments of Defense from different countries upheld serious training to mentor robots in new errands through perception [72].

### 5.4 Medical Research

Malignancy analysts are utilizing Deep figuring out how to distinguish disease cells naturally. Groups at UCLA assembled a propelled magnifying instrument that yields a high-dimensional informational index used to prepare an intelligent learning application to precisely identify malignant growth cells [73].

## 5.5 Industrial Automation

Deep learning is assisting with improving laborer wellbeing around overwhelming hardware via consequently distinguishing when individuals or items are inside a risky separation of machines [74].

## 5.6 Electronics

Deep learning is being utilized in mechanized hearing and discourse interpretation. For instance, home help gadgets that react to your voice and realize your inclinations are controlled by Deep learning applications [75].

## 5.7 Image Recognition

A general estimation set for picture characterization is that the MNIST data information set. MNIST comprises of composed digits and consists of sixty thousand training models and ten thousand investigate models. Like TIMIT, its diminutive size grants clients to check different arrangements. An extensive rundown of results is out there on this set. Deep learning-based picture acknowledgment has become "powerful," giving a ton of right outcomes than human contenders. It first occurred in 2011. Learned prepared vehicles at present decipher 360° camera sees. Another model is Facial Dysmorphology Novel Analysis (FDNA), which is utilized to examine instances of human pathology identified with a larger than average data of hereditary disorders [76].

## 5.8 Visual Craftsmanship Preparing

Firmly connected with what has happened in picture acknowledgment is that the expanding utilization of concentrated learning strategies to differed visual expressions works. For example, DNN has substantiated itself as capable [77].

- a. Distinctive the style measure of the given composition.
- b. Neural vogue Transfer—Capturing the style of a given structure and applying it to an arbitrary photo or video outwardly.
- c. Generating remarkable authentic procedures bolstered arbitrary visual info fields.

## 5.9 Natural Language Processing

Neural systems are wont to actualize language systems since the principal 2000s. LSTM improved AI and language displaying [78]. Diverse essential strategies during this field square measure negative testing and word implanting. Word inserting, like the term 2vec, is thought of as a reflective layer in an incredibly Deep learning plan that changes the

word particle into a point delineation of the concept's comparative with many words inside the dataset. The position is envisioned as a degree in a very vector house. The exploitation of the name installing as a partner degree RNN input layer allows the system to examine sentences and expressions exploitation powerful organization vector unmistakable phonetics. A build vector unmistakable etymology is thought of as an achievable setting free, engaging semantics (PCFG) implemented by partner degree RNN. Word embeddings in-assembled algorithmic auto-encoder will survey sentence closeness and find rewording. Deep neural designs give the best outcomes for body parsing, estimation investigation, information recovery, verbally expressed correspondence cognizance, AI, connecting of the reference unit, kind acknowledgment, content grouping, etc. Late improvements sum up word entering to install sentences. Google Translate (GT) utilizes start to finish long memory arrange. Google Neural AI (GNMT) uses a partner degree model-based AI technique during which the framework gains from army occurrences [79].

## 5.10 Bioinformatics

To anticipate clinical factor, claim to fame explanation and quality capacity connections, an autoencoder ANN was used in bioinformatics. In clinical data preparing, Deep learning was wont to appraise rest quality upheld information on wear and wellbeing entanglements from electronic wellbeing record information. In-depth knowledge has conjointly indicated practicality in consideration [80].

## 5.11 Medical Image Analysis

Top to bottom learning has been appeared to supply serious results like neoplastic cell order, wound identification, organ division, and picture improvement in clinical knowledge [81].

## 5.12 Mobile Publicizing

It is continuously hard to look out an excellent portable crowd for versatile promoting, as a few information guides need to be thoughtful. In this manner, the outside area used in advertisement serving is made and used by an advertisement server. Deep learning has been wonted to decipher mammoth, multi-measurement promoting datasets [82]. A few information focuses square measure gathered all through the solicitation/administration/snap of the web promoting cycle. This information will turn into AI to support a promotion decision.

### 5.13 Image Rebuilding

Deep learning has been with progress applied to switch issues like super-goals, inpainting, and film colorization. These applications encapsulate learning methodologies like "shrinkage zones for powerful picture reclamation" that train on an image dataset, and Deep Image past, that trains on an image requiring rebuilding [83].

### 5.14 Financial Misrepresentation Location

Deep learning is being with progress applied for cash misrepresentation location and tax evasion. The Deep Anti-Money wash Detection System will build up connections and similitudes among information. Accordingly, the street figures out how to spot or group oddities and anticipate explicit occasions. The appropriate response exploits each regulated learning strategies, like the arrangement of fluffy exchanges, and clueless learning, for example, oddity recognition [84].

## 6 Problems with Deep Neural Networks

As with ANNs, several problems will arise with DNNs if they're naively trained. Two common issues square measure overfitting and computation time [85, 86]. DNNs square measure in danger of over-fitting owing to extra layers of abstraction that permits them to model rare dependencies in coaching knowledge. Regularization strategies like weight reduction (L2-regularization) or meagerness (L-1-regularization) are applied to coaching to assist combat overfitting. Another recent regularization methodology used to DNNs is dropout regularization. In dropouts, some variety of unit's square measure indiscriminately born from hidden layers throughout coaching. This helps break down the rare dependencies which will occur in coaching knowledge [87]. The effective methodology for coaching these structures is error-correction coaching (such as backpropagation on gradient descent), because of its simple implementation and its tendency to convert to raised native optima than different coaching strategies. However, these strategies are computationally exclusive, particularly for Deep Neural Networks. There square measure many coaching parameters to contemplate with DNN, like size (number of layers and variety of units per layer), learning rate, and initial weights. Intensive through parameter houses might not be attainable for best parameters because of value in time and procedure resources. Numerous 'tricks' (calculating gradients on multiple coaching instances quickly rather than individual cases) victimization mini batching are shown to hurry

up the calculation [88]. The massive process turnout of graphics process units (GPUs) because of matrix and vector computation has created vital speedups in coaching, that should be GPU-friendly [89]. Radical alternatives to the bench like extreme learning machines, "no-prop" networks, coaching perennial networks while not backtracking, and weightless networks square measure attracting attention.

### 6.1 Data Labeling

Most current computer science models square measure trained through supervised learning. This implies that humans should label and classify the underlying knowledge, which might be an oversized and fallible core. For example, corporations rising self-driving-car technology square measure hiring many individuals to annotate hours of the video manually feeds from model vehicles to assist train these systems [90].

### 6.2 Obtain Massive Training Datasets

It is found that easy deep learning techniques like CNN, in some cases, mimic the information of specialists in medication and different fields. However, this wave of machine learning needs coaching knowledge sets that aren't solely labeled, conjointly sufficiently comprehensive and universal. Intensive learning strategies were required for legion individuals to become comparatively sensible at classification tasks and, in some cases, to perform at the number of humans. While no surprise, deep learning is well-known among large technical school corporations. They are victimization extensive knowledge to accumulate petabytes of information. Which permits them to make associate degree cogent and correct learning model [91].

### 6.3 Automatic Colorization of Black and White Images

Picture colorization is the issue of adding concealing to high differentiation photos. Customarily this was done by hand with the human effort since it is such a problematic assignment. Profound learning can be used to use the things, and their setting inside the photograph to concealing the image, much like a human overseer, may advance toward the issue. A visual and outstandingly incredible accomplishment is made. Such capacity impacts the enormous convolutional neural frameworks which are arranged for ImageNet and co-settled on the topic of picture colorization. For the most part, the strategy incorporated the use of large convolutional neural systems and managed layers that recreate the image with the alternative of shading [92].

## 6.4 Consequently, Adding Sounds to Silent Movies

At present, the system must incorporate sounds to arrange a peaceful video. The structure is set up to use 1000 occasions of video with the sound of a drumstick striking different surfaces and making various sounds. A Deep learning model accomplice the video diagrams with a database of pre-rerecorded sounds to pick a sound to play that best matches what's happening in the scene. The system is using a Turing-test like a course of action where individuals expected to make sense of which video had the certifiable or the fake (organized) sounds. New use of both convolutional neural frameworks and LSTM discontinuous neural frameworks [93].

## 6.5 Object Classification and Detection in Photographs

This endeavour requires the game plan of things inside a photograph as one of a ton of as of late known articles. Top tier results have been cultivated on benchmark examples of this issue using huge convolutional neural frameworks. An achievement right now Alex Krizhevsky et al. [94] results in the ImageNet arrangement issue called AlexNet.

## 6.6 Programmed Image Caption Generation

Modified picture subtitling is the place given an image, and the system must deliver an engraving that depicts the substance of the image [95]. In 2014, there was an impact of Deep learning counts achieving incredibly essential results on this issue, using the work from top models for object request and thing revelation in photos. When you can recognize dissents in photographs and make marks for those articles, you can see that the accompanying stage is to change those names into a level-headed sentence portrayal. This is one of those results that blew my mind and still does. Extraordinarily incredible for sure. For the most part, the structures incorporate the usage of enormous convolutional neural frameworks for the thing area in the photographs and, after that, an irregular neural system like an LSTM to change the names into a mindful sentence [96].

## 6.7 Programmed Handwriting Generation

It is the place given a corpus of handwriting models, produce new handwriting for a given the word or expression [97]. The content is given as a gathering of headings used by a pen when the handwriting tests were made. From this corpus, the association between the pen advancement and the letters is discovered, and new models can be made specially appointed. It is charming that different styles can be scholarly, and a while later reflected. I was unable to need

anything over to see this work got together with some criminological hand forming assessment inclination [98].

## 6.8 Programmed Text Generation

It is an exciting endeavour, where a corpus of substance is discovered. From this model, a new content is made, word-by-word, or character-by-character [99]. The model is fit for making sense of how to spell, emphasize, structure sentences, and even catch the style of the substance in the corpus. Enormous discontinuous neural frameworks are used to get acquainted with the association between things in the groupings of data strings and a while later produce content. More starting late LSTM intermittent neural frameworks are displaying unbelievable achievement on this issue using a character-based model, making each character in turn [100].

## 7 Case Study: Handwriting Recognition Using Deep Learning

To feature the utilization of profound learning in different applications, we center around the usage of feedforward neural systems and Convolutional Neural Networks on written by hand digits in the MNIST database. The MNIST database comprises of pictures of sizes  $28 \times 28$ . For the feedforward neural system, the lattice is reshaped to  $784 \times 1$  and took care of as a contribution to the system. We reshape the picture network into a vector of size  $784 \times 1$  and feed it as information. Accordingly, 60,000 models are prepared to utilize these systems, and 10,000 examples are tried to perform the characterization of digits. The engineering utilized for both feedforward neural systems just as Convolutional Neural Networks appears in the tables.

The accuracy of the order of digits utilizing feedforward neural systems after 20 emphases is 93.76 percent. While with the utilization of Convolutional neural networks, the precision goes up to 99.2% and this has been arranged in the table underneath:

Layer	Number of neurons
Input layer	784
Hidden layer 1	256
Hidden layer 2	256
Output layer	10

Layer	Number and size of filters	
Convolution layer	$32, 3 \times 3$	ReLU
Maxpooling	$1, 2 \times 2$	
Convolution layer	$64, 3 \times 3$	ReLU



Layer	Number and size of filters	
Maxpooling	1,2×2	
Fully connected	128	ReLu
Fully connected	10	SoftMax

Iterations	Type of network	Accuracy achieved (%)
20	Feedforward Neural nets	93.76
20	CNN	99.20

## 8 Conclusion

Deep learning is genuinely a speedily expanding application of machine learning. Many applications related to it delineated higher than prove its fast development in only some years. The employment of those algorithms in numerous fields demonstrates its skillfulness. The publication analysis performed during this study reflects the relevancy of this method and reflects the tendency towards deep learning and development of future analysis during this space. Moreover, it's vital to notice that hierarchies of layers and direction in learning square measure essential features to improve a roaring application concerning deep learning. The regime is necessary for correct knowledge classification, whereas guideline considers the importance of the information as a part of the method. The worthiest of deep learning on the improvement of existing applications in machine learning is because of its novelty on hierarchic layer process. Deep learning will provide effective results on the digital image process and speech recognition. The reduction in error share (10 to 20%) confirms the advance compared to the present and tested strategies. Throughout this era and within the future, deep learning may result in a beneficial security tool because of face recognition and speech recognition. Additionally, the digital image process could be an analysis space that will be applied in several fields. For this reason and when proving to be a real adaptation, in-depth information could be a new and exciting subject of advancement in computer science.

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Fourie C (2003) Deep learning? What deep learning? South Afr J High Educ. <https://doi.org/10.4314/sajhe.v17i1.25201>
- Pandey H (2017) Genetic algorithm for grammar induction and rules verification through a PDA simulator. IAES Int J Artif Intell (IJ-AI) 6(3):100. <https://doi.org/10.11591/ijai.v6.i3.pp100-111>
- Ordóñez C, Zhang Y, Johnson S (2018) Scalable machine learning computing a data summarization matrix with a parallel array DBMS. Distrib Parallel Databases 37(3):329–350. <https://doi.org/10.1007/s10619-018-7229-1>
- Criminisi A (2016) Machine learning for medical images analysis. Med Image Anal 33:91–93. <https://doi.org/10.1016/j.media.2016.06.002>
- Goel A (2017) Editorial: expository AI applications. AI Mag 38(1):3. <https://doi.org/10.1609/aimag.v38i1.2739>
- Alonso E (2002) AI and agents: state of the art. AI Mag 23(3):25–25
- Shapiro, S. C. (1971). The MIND system: a data structure for semantic information processing (Vol. 837, No. PR). Rand Corp Santa Monica Califss
- Beck J, Stern M, Haugsjaa E (1996) Applications of AI in education. XRDS Crossroads ACM Mag Stud 3(1):11–15
- Faught WS (1986) Applications of AI in engineering. Computer 7:17–27
- Zhenan S, Zhaoxiang Z, Wei W, Fei L, Tieniu T (2020) Artificial intelligence: developments and advances in 2019. Front Data Comput 1(2):1–16
- Li XH, Cao CC, Shi Y, Bai W, Gao H, Qiu L, Chen L (2020) A survey of data-driven and Knowledge-aware eXplainable AI. IEEE Trans Knowl Data Eng. <https://doi.org/10.1109/TKDE.2020.2983930>
- Maxwell AE, Warner TA, Fang F (2018) Implementation of machine-learning classification in remote sensing: an applied review. Int J Remote Sens 39(9):2784–2817
- Usama M, Qadir J, Raza A, Arif H, Yau KLA, Elkhatib Y, Al-Fuqaha A (2019) Unsupervised machine learning for networking: techniques, applications and research challenges. IEEE Access 7:65579–65615
- Osband, I., Doron, Y., Hessel, M., Aslanides, J., Sezener, E., Saraiva, A., & Van Roy, B. (2019). Behaviour suite for reinforcement learning. arXiv preprint
- Mnih, V., Badia, A. P., Graves, A. B., Harley, T.J. A., Silver, D., & Kavukcuoglu, K. (2019). U.S. Patent Application No. 16/403,388.
- Mnih, V., Czarnecki, W., Jaderberg, M. E., Schaul, T., Silver, D., & Kavukcuoglu, K. (2019). U.S. Patent Application No. 16/403,385.
- Mnih, V., Czarnecki, W., Jaderberg, M. E., Schaul, T., Silver, D., & Kavukcuoglu, K. (2019). U.S. Patent Application No. 16/403,385
- Zhai, X., Oliver, A., Kolesnikov, A., & Beyer, L. (2019). S4I: self-supervised semi-supervised learning. In: proceedings of the IEEE international conference on computer vision (pp. 1476–1485)
- Souza JTD, Francisco ACD, Piekarski CM, Prado GFD (2019) Data mining and machine learning to promote smart cities: a systematic review from 2000 to 2018. Sustainability 11(4):1077
- Niculescu V (2019) High performance computing in big data analytics. Appl Med Inform 41:2–21
- Zhao ZQ, Zheng P, Xu ST, Wu X (2019) Object detection with deep learning: a review. IEEE Trans Neural Netw Learn Syst 30(11):3212–3232
- Huang, S. L., Xu, X., Zheng, L., & Wornell, G. W. (2019). An information theoretic interpretation to deep neural networks.



- In: 2019 IEEE international symposium on information theory (ISIT) (pp. 1984–1988). IEEE
23. Elfving S, Uchibe E, Doya K (2018) Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural Netw* 107:3–11
  24. Szegedy, C., Erhan, D., & Toshev, A. T. (2016). U.S. Patent No. 9,275,308. Washington, DC: U.S. Patent and Trademark Office
  25. Kombrink, Stefan & Mikolov, Tomas & Karafiát, Martin & Burget, Lukas. (2011). Recurrent neural network based language modeling in meeting recognition. In: proceedings of the annual conference of the international speech communication association, INTERSPEECH. 2877–2880
  26. Hinton G (2012) Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process Mag* 29(6):82–97
  27. Chung, J., Ahn, S., & Bengio, Y. (2016). Hierarchical multiscale recurrent neural networks. arXiv preprint
  28. Li, B., & Sainath, T. N. (2018). U.S. Patent No. 9,984,683. Washington, DC: U.S. Patent and Trademark Office
  29. Martin, L. J., Ammanabrolu, P., Wang, X., Hancock, W., Singh, S., Harrison, B., & Riedl, M.O. (2018). Event representations for automated story generation with deep neuralnets. In: thirty-second AAAI conference on artificial intelligence
  30. Ren S, He K, Girshick R, Zhang X, Sun J (2016) Object detection networks on convolutional feature maps. *IEEE Trans Pattern Anal Mach Intell* 39(7):1476–1481
  31. Miyato T, Maeda SI, Koyama M, Ishii S (2018) Virtual adversarial training: a regularization method for supervised and semi-supervised learning. *IEEE Trans Pattern Anal Mach Intell* 41(8):1979–1993
  32. Wang M, Fu W, Hao S, Tao D, Wu X (2016) Scalable semi-supervised learning by efficient anchor graph regularization. *IEEE Trans Knowl Data Eng* 28(7):1864–1877
  33. Gong C, Tao D, Maybank SJ, Liu W, Kang G, Yang J (2016) Multi-modal curriculum learning for semi-supervised image classification. *IEEE Trans Image Process* 25(7):3249–3260
  34. Tang, H., Xu, D., Yan, Y., Corso, J. J., Torr, P. H., & Sebe, N. (2020). Multi-channel attention selection GANs for guided image-to-image translation. arXiv preprint
  35. Royer A, Bousmalis K, Gouws S, Bertsch F, Mosseri I, Cole F, Murphy K (2020) Xgan: unsupervised image-to-image translation for many-to-many mappings. In *Domain adaptation for visual understanding*. Springer, Cham, pp. 33–49
  36. French, G., Mackiewicz, M., & Fisher, M. (2017). Self-ensembling for visual domain adaptation. arXiv preprint
  37. Saito, K., Ushiku, Y., & Harada, T. (2017). Asymmetric tri-training for unsupervised domain adaptation. In: *proceedings of the 34th international conference on machine learning-volume 70* (pp. 2988–2997). JMLR. org
  38. Hung, W. C., Tsai, Y. H., Liou, Y. T., Lin, Y. Y., & Yang, M. H. (2018). Adversarial learning for semi-supervised semantic segmentation. arXiv preprint
  39. Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double q-learning. In: thirtieth AAAI conference on artificial intelligence
  40. M. Hausknecht and P. Stone (2015). Deep recurrent q-learning for partially observable mdps. AAAI
  41. Hessel, M., Modayil, J., Van Hasselt, H., Schaul, T., Ostrovski, G., Dabney, W., & Silver, D. (2018). Rainbow: combining improvements in deep reinforcement learning. In thirty-second AAAI conference on artificial intelligence
  42. Wu, Y., Mansimov, E., Grosse, R. B., Liao, S., & Ba, J. (2017). Scalable trust-region method for deep reinforcement learning using kronecker-factored approximation. In: *advances in neural information processing systems* (pp. 5279–5288)
  43. Dabney, W., Rowland, M., Bellemare, M. G., & Munos, R. (2018, April). Distributional reinforcement learning with quantile regression. In: thirty-second AAAI conference on artificial intelligence
  44. Marcus, G. (2018). Deep learning: a critical appraisal. arXiv preprint
  45. Shu, M. (2019). Deep learning for image classification on very small datasets using transfer learning
  46. Jin L, Li S, La HM, Luo X (2017) Manipulability optimization of redundant manipulators using dynamic neural networks. *IEEE Trans Ind Electron* 64(6):4710–4720
  47. Bertinetto, L., Henriques, J. F., Valmadre, J., Torr, P., & Vedaldi, A. (2016). Learning feed-forward one-shot learners. In: *advances in neural information processing systems* (pp. 523–531)
  48. SaishanmugaRaja V, Rajagopalan SP (2013) IRIS recognition system using neural network and genetic algorithm. *Int J Comput Appl* 68(20):49–53. <https://doi.org/10.5120/11699-7431>
  49. Tran, D. T., Kannianen, J., Gabbouj, M., & Iosifidis, A. (2019). Data-driven neural architecture learning for financial time-series forecasting. arXiv preprint
  50. Qi H, Shi Y, Tian Y, Mayhew C, Yu DL, Gomm JB, Zhang Q (2019) A new fault diagnosis and fault-tolerant control method for mechanical and aeronautical systems with neural estimators. *Adv Mech Eng* 11(11):1687814019891659
  51. Huang, Y., Zhang, Y., & Yan, C. (2019). Automatic sleep staging based on deep neural network using single channel EEG. In: *international conference on knowledge management in organizations* (pp. 63–73). Springer: Cham
  52. Martens J, & Sutskever I (2011). Learning recurrent neural networks with hessian-free optimization. In: *proceedings of the 28th international conference on machine learning (ICML-11)* (pp. 1033–1040)
  53. Tolosana R, Vera-Rodriguez R, Fierrez J, Ortega-Garcia J (2018) Exploring recurrent neural networks for on-line handwritten signature biometrics. *IEEE Access* 6:5128–5138
  54. Kasongo SM, Sun Y (2019) A deep long short-term memory based classifier for wireless intrusion detection system. *ICT Express* 6(2):98–103
  55. Bouktif S, Fiaz A, Ouni A, Serhani MA (2020) Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting. *Energies* 13(2):391
  56. Zhang YF, Fitch P, Thorburn PJ (2020) Predicting the trend of dissolved oxygen based on the kPCA-RNN model. *Water* 12(2):585
  57. Razzak, F., Yi, F., Yang, Y., & Xiong, H. (2019). An integrated multimodal attention-based approach for bank stress test prediction. In: 2019 IEEE international conference on data mining (ICDM) (pp. 1282–1287). IEEE
  58. Kumar, P., Gupta, M., Gupta, M., & Sharma, A. (2019). Profession Identification Using Handwritten Text Images. In: *international conference on computer vision and image processing* (pp. 25–35). Springer: Singapore
  59. Yan LC, Bengio YS, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
  60. Tong W, Li L, Zhou X, Hamilton A, Zhang K (2019) Deep learning PM 2.5 concentrations with bidirectional LSTM RNN. *Air Qual Atmos Health* 12(4):411–423
  61. Vedaldi A, & Lenc K (2015). Matconvnet: Convolutional neural networks for matlab. In: *proceedings of the 23rd ACM international conference on multimedia* (pp. 689–692)
  62. Aloysius N, Geetha M (2017). A review on deep convolutional neural networks. In: *2017 international conference on communication and signal processing (ICCSP)* (pp. 0588–0592). IEEE
  63. Chi YS, Kamarulzaman SF (2020) Intelligent gender recognition system for classification of gender in malaysian

- demographic. In: Nasir ANK, Ahmad MA, Najib MS (eds) In ECCE 2019. Springer, Singapore, pp 283–295
64. Wang C, Zhang Z, Zhou X (2018) An image copy-move forgery detection scheme based on A-KAZE and SURF features. *Symmetry* 10(12):706
  65. Gan W, Wang S, Lei X, Lee MS, Kuo CCJ (2018) Online CNN-based multiple object tracking with enhanced model updates and identity association. *Signal Process Image Commun* 66:95–102
  66. Kong X, Gong S, Su L, Howard N, Kong Y (2018) Automatic detection of acromegaly from facial photographs using machine learning methods. *EBioMedicine* 27:94–102
  67. Telgarsky, M. (2015). Representation benefits of deep feedforward networks. arXiv preprint
  68. Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE (2017) A survey of deep neural network architectures and their applications. *Neurocomputing* 234:11–26
  69. Yan, R., Song, Y., & Wu, H. (2016) Learning to respond with deep neural networks for retrieval-based human-computer conversation system. In: proceedings of the 39th international ACM SIGIR conference on research and development in information retrieval (pp. 55–64)
  70. Hatcher WG, Yu W (2018) A survey of deep learning: platforms, applications and emerging research trends. *IEEE Access* 6:24411–24432
  71. Milz, S., Arbeiter, G., Witt, C., Abdallah, B., & Yogamani, S. (2018). Visual slam for automated driving: Exploring the applications of deep learning. In: proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 247–257)
  72. Najafabadi MM, Villanustre F, Khoshgoftaar TM, Seliya N, Wald R, Muharemagic E (2015) Deep learning applications and challenges in big data analytics. *J Big Data* 2(1):1
  73. Shen D, Wu G, Suk HI (2017) Deep learning in medical image analysis. *Annu Rev Biomed Eng* 19:221–248
  74. Zhang K, Zhu Y, Leng S, He Y, Maharjan S, Zhang Y (2019) Deep learning empowered task offloading for mobile edge computing in urban informatics. *IEEE Internet Things J* 6(5):7635–7647
  75. Lemley J, Bazrafkan S, Corcoran P (2017) Deep learning for consumer devices and services: pushing the limits for machine learning, artificial intelligence, and computer vision. *IEEE Consum Electron Mag* 6(2):48–56
  76. Wu, R., Yan, S., Shan, Y., Dang, Q., & Sun, G. (2015). Deep image: Scaling up image recognition. arXiv preprint
  77. Noda K, Yamaguchi Y, Nakadai K, Okuno HG, Ogata T (2015) Audio-visual speech recognition using deep learning. *Appl Intell* 42(4):722–737
  78. Deng L, Liu Y (eds) (2018) Deep learning in natural language processing. Springer, Singapore
  79. Sun S, Luo C, Chen J (2017) A review of natural language processing techniques for opinion mining systems. *Inf Fusion* 36:10–25
  80. Min S, Lee B, Yoon S (2017) Deep learning in bioinformatics. *Brief Bioinform* 18(5):851–869
  81. Ker J, Wang L, Rao J, Lim T (2017) Deep learning applications in medical image analysis. *IEEE Access* 6:9375–9389
  82. Nakao A, Du P (2018) Toward in-network deep machine learning for identifying mobile applications and enabling application specific network slicing. *IEICE Trans Commun.* <https://doi.org/10.1587/transcom.2017CQI0002>
  83. Kumar, D., Kumar, C., & Shao, M. (2017). Cross-database mammographic image analysis through unsupervised domain adaptation. In: 2017 IEEE international conference on big data (Big Data) (pp. 4035–4042). IEEE
  84. Wang Y, Xu W (2018) Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. *Decis Support Syst* 105:87–95
  85. Zhang, C., Vinyals, O., Munos, R., & Bengio, S. (2018). A study on overfitting in deep reinforcement learning. arXiv preprint
  86. Graves, A. (2016). Adaptive computation time for recurrent neural networks. arXiv preprint
  87. Molchanov, D., Ashukha, A., & Vetrov, D. (2017). Variational dropout sparsifies deep neural networks. In: *proceedings of the 34th international conference on machine learning-volume 70*: (pp. 2498–2507). JMLR. org
  88. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. arXiv preprint
  89. Cui, H., Zhang, H., Ganger, G. R., Gibbons, P. B., & Xing, E. P. (2016, April). Geeps: Scalable deep learning on distributed gpus with a gpu-specialized parameter server. In: proceedings of the eleventh European conference on computer systems (pp. 1–16)
  90. Wang, D., & Shang, Y. (2014). A new active labeling method for deep learning. In: 2014 international joint conference on neural networks (IJCNN) (pp. 112–119). IEEE
  91. Xiao, T., Xia, T., Yang, Y., Huang, C., & Wang, X. (2015). Learning from massive noisy labeled data for image classification. In: proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2691–2699)
  92. Larsson, G., Maire, M., & Shakhnarovich, G. (2016). Learning representations for automatic colorization. In: European conference on computer vision (pp. 577–593). Springer: Cham
  93. Ghose, S., & Prevost, J. J. (2020). AutoFoley: artificial synthesis of synchronized soundtracks for silent videos with deep learning. arXiv preprint
  94. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 60(6):1097–1105
  95. Bai S, An S (2018) A survey on automatic image caption generation. *Neurocomputing* 311:291–304
  96. Hossain MZ, Sohel F, Shiratuddin MF, Laga H (2019) A comprehensive survey of deep learning for image captioning. *ACM Comput Surv (CSUR)* 51(6):1–36
  97. Wang, H., Qin, Z., & Wan, T. (2018). Text generation based on generative adversarial nets with latent variables. In: Pacific-Asia conference on knowledge discovery and data mining (pp. 92–103). Springer: Cham
  98. Jaderberg M, Simonyan K, Vedaldi A, Zisserman A (2016) Reading text in the wild with convolutional neural networks. *Int J Comput Vision* 116(1):1–20
  99. Guo, J., Lu, S., Cai, H., Zhang, W., Yu, Y., & Wang, J. (2018). Long text generation via adversarial training with leaked information. In: thirty-second AAAI conference on artificial intelligence
  100. Wiseman, S., Shieber, S. M., & Rush, A. M. (2017). Challenges in data-to-document generation. arXiv preprint

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.