International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 215-218 www.ijaem.net ISSN: 2395-5252

"Study on Computer Vision Approaches based on Deep learning and Neural networks".

Khushboo Takleguide : Prof.Mutkule.P.R

Submitted: 10-06-2021 Revised: 23-06-2021 Accepted: 26-06-2021

ABSTRACT
It's nice to realize that we humans can manage to enable machines with our own natural skills: learning by example and perceiving the outside world. But the problem is that significantly more time and effort is needed to teach computers how to "see" like we humans do. But if we think for practical purpose this capability already brings to Organizations and Businesses, the effort pays off.
Over the years DL methods have been outperform previous state of the art machine learning techniques in several fields, with computer vision being one of the most prominent cases around.The studyindicates that the field is a promising area for research.
The purpose of this paper is to carry outa scientific mapping observe, to analyzepresentstudiesapproximately implementations of laptopimaginative and prescienttacticsprimarily based totally on deep studying algorithms.

I. INTRODUCTION

DL allows computational models of multiple processing layers to learn and represent data with multiple levels of abstraction showing how the brain perceives and understands multimodal information, thus implicitly capturing intricate structures of large-scale data. Deep learning is a rich family of methods, encompassing neural networks, hierarchical probabilistic models, and a variety of unsupervised and supervised feature learning algorithms.

Deep learning is based Computer vision that enables computers to understand and extract the information form the content of images. The goal of computer vision is to automate tasks that the human vision system can do with an proper accuracy and higher speed. It has its great strides in a variety of computer vision problems, such as object detectionmotion tracking, action recognition, human pose estimation and semantic segmentation.

The study indicates that this field is a promising area for research.

The research in computer vision was started way back in 1950s. Since then, it has come a long way but it still finds us far from the ultimate objective. But with neural networks and deep learning, the study has become empowered like never before. In this paper, I will also introduce you to Convolution Neural Networks which form the crux of DL applications in computer vision.

II. DEEP LEARNING METHODS AND DEVELOPMENTS 2.1. CONVOLUTIONAL NEURAL NETWORKS.

Convolutional Neural Networks (CNNs) were inspired by the visual system's structure, and in particular by the models of it proposed in . The first computational models based on these local connectivity's between neurons and on hierarchically organized transformations of the image are found in Neocognitron , which describes that when neurons with the same parameters are applied on patches of the previous layer at different locations, a form of translational invariance is acquired. Yann LeCun and his colleagues later designed Convolutional Neural Networks employing the error gradient and attaining very good results in a variety of pattern recognition tasks. A CNN mainly have three types of neural layers, namely, **(i) convolutional layers, (ii) pooling layers, and (iii) fully connected layers.** Each type of layer has a different role. Figure 2.1.1 shows CNN architecture for object detection in image task. Every layer of a CNN transforms the input volume to an output volume of neuron activation, eventually leading to the final fully connected layers, resulting in a mapping of the input data to a 1D feature vector. CNNs have been extremely successful in computer vision applications, such as face recognition, object detection, powering vision in robotics, and self-driving cars.

(i) Convolutional Layers.
In the convolutional layers, CNN utilizes various kernels to convolve the whole image as

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 215-218 www.ijaem.net ISSN: 2395-5252

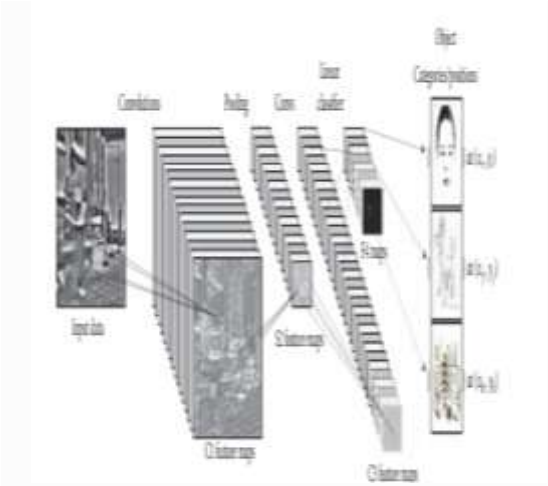

Fig 2.1.1 : Example architecture of a CNN for computer vision task (Object Detection).

Well as the intermediate feature maps, generating various feature maps. Because of the advantages of the convolution operation, several works have proposed it as a substitute for fully connected layers with a view to attaining faster learning times.

(ii) Pooling Layers.

Pooling layers are in charge of reducing the spatial dimensions (width × height) of the input volume for the next convolutional layer. The pooling layer does not affect the depth dimension of the volume. The operation performed by this layer is also called subsampling or down sampling, as the reduction of size leads to a simultaneous loss of information. However, such loss is beneficial for the network because the decrease in size leads to less computational overhead for the upcoming layers of the network, and also it works against over fitting. Average pooling and max pooling are the most commonly used strategies. In detailed theoretical analysis of max pooling and average pooling performances is given, whereas in it was shown that max pooling can lead to faster convergence, select superior invariant features, and improve generalization. Also there are other variations of the pooling layer in the literature, each inspired by different motivations and serving distinct needs, for example, stochastic pooling , spatial pyramid pooling , and def-pooling .

(iii) Fully Connected Layers.

Following several convolutional and pooling layers, the high-level reasoning in the neural network is performed via fully connected layers. Neurons in a fully connected layer have full connections to all activation in the previous layer, as their name implies. Their activation can hence be computed with a matrix multiplication followed by a bias offset. Fully connected layers eventually convert the 2D feature maps into a 1D feature vector. The derived vector either could be fed forward into a certain number of categories for classification or could be considered as a feature vector for further processing. The architecture of CNNs employs three concrete ideas: (a) local receptive fields, (b) tied weights, and (c) spatial subsampling. Based on local receptive field, each unit in a convolutional layer receives inputs from a set of neighboring units belonging to the previous layer. Tis way neurons are capable of extracting elementary visual features such as edges or corners. These features are then combined by the subsequent convolutional layers in order to detect higher order features. Furthermore, the idea that elementary feature detectors, which are useful on a part of an image, are likely to be useful across the entire image is implemented by the concept of tied weights. The concept of tied weights constraints a set of units to have identical weights. Concretely, the units of a convolutional layer are organized in planes. All units of a plane share the same set of weights. Thus, each plane is responsible for constructing a specific feature. Te outputs of planes are called feature maps. Each convolutional layer consists of several planes, so that multiple feature maps can be constructed at each location.

2.2 Generative Adversarial Networks :

GAN is a combination of two **deep learning techniques** of neural networks – a Generator and a Discriminator. While the Generator Network yields artificial data, the Discriminator helps in discerning between a real and a false data.

Both of the networks are competitive, as the Generator keeps producing artificial data identical to real data – and the Discriminator continuously detecting real and unreal data. In a scenario where there's a requirement to create an image library, the Generator network would produce simulated data to the authentic images. It would then generate a deconvolution neural network.

It would then be followed by an Image Detector network to differentiate between the real and fake images. Starting with a **50% accuracy chance**, the detector needs to develop its quality of classification since the generator would grow better in its artificial image generation. Such competition would overall contribute to the network in its effectiveness and speed.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 7 July 2021, pp: 215-218 www.ijaem.net ISSN: 2395-5252

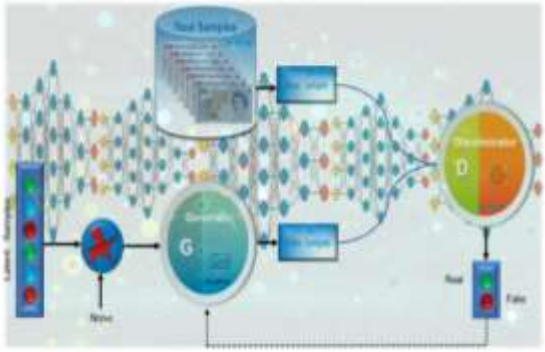

Fig 2.2.1 :Example for Generative Adversarial Networks

2.3 Multilayer Perceptrons (MLPs)

MLPs are an excellent place to start learning about deep learning technology. MLPs belong to the class of feedforward neural networks with multiple layers of perceptrons that have activation functions. MLPs consist of an input layer and an output layer that are fully connected. It has the same number of input and output layers but may have multiple hidden layers and can be used to build speech-recognition, image-recognition, and machine-translation software.
Below is an example of an MLP.

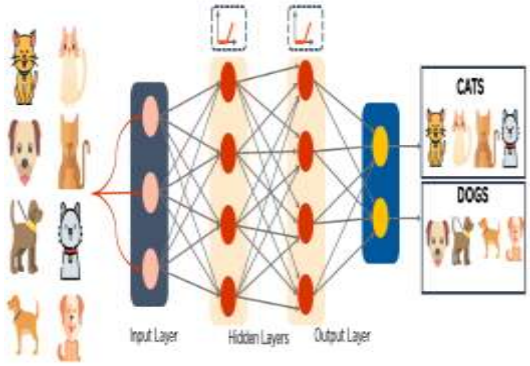

Fig 2.3.1: The diagram shows weights and bias and applies suitable activation functions to classify images of cats and dogs.

III. LITERATURE SURVEY

Deep learning (DL) is an important machine learning field that has achieved considerable success in many research areas. In the last decade, the-state-of-the-art studies on many research areas such as computer vision, object recognition, speech recognition, and natural language processing were especially led to the awakening of the artificial intelligence from deep sleep. Nowadays, many researchers try to find solutions to many problems in various fields under the light of DL methods.

In this study, it is presented important knowledge to guide about DL models and challenging topics that can be used in DL for researchers. This study investigated DL studies which are made in the most popular and challenging fields such as autonomous vehicles, natural language processing, handwritten character recognition, signature verification, voice and video recognition, medical image processing, and big data.

Furthermore, this study points out the remaining challenges of these research areas that can be solved by DL, and discusses future topics to help the researchers. In the present day, Deep learning methods have reached better results than humans in object recognition. According to the literature studies on DL, It is foreseen that this success will be achieved in areas such as autonomous vehicles, medical image processing, big data analysis, and character recognition.

IV. MOTIVATION

DL is the most significant development in the field of computer science in recent times.Its effect has been felt in almost all medical fields.It is already disrupting and transforming businesses and industries. There is a race among the world's leading economies and technology companies to advance deep learning.There are already many regionsin which deep studying has passed human degreefunctionality and overall performance.

This paper covers the different types of deep learning network architectures, deep learning algorithms, their shortcomings, optimization methods and the latest implementations and applications.

V. FUTURE SCOPE FOR DEEP LEARNING

Deep learning has a varied range of applications, which has led to a rise in its popularity and its usage in various industries. It is used by several organizations from different sectors or industries. Some fields of application of deep learning are :

- **Image and fingerprint recognition functions**
- **Banking apps**
- **Medical research tool**
- **Business trends and outcomes**

Deep learning algorithms seek to explore and analyze the unknown structure in the input distribution to discover useful representations with the help of feature learning, continually evolving and improvising with the input of additional data.

The scope of DL knowledge for its self-adaptive feature is boundless in today's time. The deep learning discipline isn't simple, and its complete capabilities have not yet been explored. But it is a potent branch of machine learning and has a lot of scopes soon.

- **It will adopt a core set of standards and tooling frameworks.**
- **It is destined to be endorsed and demanded by several brands and businesses.**
- **Fast coding through simplified programming frameworks.**
- **It will be increasingly used in the field of designing and building generative designing tools.**
- **Image search, photorealistic rendering, resolution enhancement, and various other applications of deep learning is being developed and researched for best results.**

Deep learning has been one of the most dynamic and versatile branches of data science that is ever-transforming and has immense potential. It stands to be one of the most promising career paths with a diverse range of scopes and opportunities.

VI. CONCLUSION

The surge of deep getting to knowduring the last years is to a first rateamountdue to the strides it has enabled with inside the place of computervision. It givesandefine of country-of-the paintingsresearchwith inside the place. It paves the way for in addition have a take a study of video evaluation, an area now not tackled as loads as imagesthrumanner of way of computervision community. CNNs have the perfectcapability of functiongetting to know, this is, of routinelygetting to knowtalentsbased totallydefinitelyon the given dataset. CNNs are also invariant to ameliorations, it is a first rate asset for fantasticcomputervision packages. Though deep neural networks gainaccuratenormaloverall performance on many duties, they however have many houses that need to be investigated and justified. The fulfillment of deep getting to knowstrategies is developing with new techniques evolved. Nowadays, DL achieves betterfulfillment than humans in masses of areaswhich includeobjectreputation. Nowadays, DL achieves higherachievement than peoplein lots ofregionsinclusive ofitempopularity.

REFERENCES

[1]. http://www.cs.toronto.edu/~ranzato/publications/cfp_ijcv_si_deeplearning.pdf.
[2]. https://jameskle.com/writes/computer-vision
[3]. https://ieeexplore.ieee.org/document/7977207/
[4]. https://pubmed.ncbi.nlm.nih.gov/29487619/
[5]. https://www.researchgate.net/profile/Ajay-Shrestha/publication/332662087_Review_of_Deep_Learning_Algorithms_and_Architectures/links/5e09f5ffa6fdcc283749fc47/Review-of-Deep-Learning-Algorithms-and-Architectures.pdf
[6]. https://startuptown.wordpress.com/2020/08/01/elon-musk-is-wrong-again/
[7]. https://arxiv.org/pdf/1805.04025
[8]. https://www.analyticsinsight.net/top-10-computer-vision-tools-for-2020/