International Journal For Technological Research In Engineering
Volume 8, Issue 5, January-2021 *(Online): 2347 - 4718*

COMPUTER VISION & IMAGE PROCESSING

[1]Devendra Joshi, [2]Nipun Jindal, [3]Gurpreet kaur
[1,2]Student, [3]Assistant Professor
Department of Computer Science & Engineering
BMCEM, Sonipat, India

Abstract: Computer vision has been studied for many years now and has developed a lot. It includes the transformation of raw data recording into techniques and digital image processing, pattern recognitions, machine learning and computer graphics. Computer vision helps scholars and researchers to analyze images and video to obtain necessary results and to perform their research accordingly. We have analyzed different developments in this field and have characterized computer vision into four sub parts e.g. image processing, object recognition, and machine learning. We have provided a brief explanation of the introductory part of Computer Vision. This paper is limited to Image processing.

Keywords: Computer Vision, Image Processing, object Recognition

1. Introduction

Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do.[1] Computer vision also focuses on automatic extraction, analysis and understanding of all the possible details form the sample image provided to the program.

In the 1960's, Computer Vision started in universities which were leading and involving new ideas in Artificial Intelligence. At that time it was meant to mimic human vision, as a beginner block to establishing intelligence in robots. The main thing that separated computer vision from other projects and researches in digital image processing was the idea of capturing a 3-dimensional image and with the help of it creating a full scene.

In the 1970's all the early algorithms regarding computer vision were designed and are used till now. Some examples are labeling of lines, optical flow, motion estimation and many more.

In the next decades this field became more and more notable and a lot of research started in this section such as snakes by Michael Kass, Andrew Witkin, and Demetri Terzopoulos. During this time one more important thing about computer vision was solved and that was treating mathematical factors within the same framework as regularization and Markov random fields. Some of the old topics also became active in this duration such as 3-D reconstructions, many more. In this time the first ever use of statistical learning techniques were done in order to practice recognition of face in image. In the end of 1990 one of the important changes was the increase in interaction between the fields of computer graphics and computer vision.

2. APPLICATIONS

The following are some of the applications of Computer vision:

1) Medicine: In this field, computer vision is used in detection of tumour, arteriosclerosis or other malign change. It also supports medical research by providing new information regarding their field with high accuracy.
2) Machine vision: The application of computer vision in industry is also known as Machine Vision, in this information is gathered to increase the manufacturing process and also in detection of faulty products during production.
3) Military: One of the increasing uses of computer vision is in the military. It is used in drones to do surveillance and also in all kinds of special missions that require full accuracy.

3. IMAGE PROCESSING

The classical problem in computer vision, image processing, and machine vision is that of determining whether or not the image data contains some specific object, feature, or activity.[1] Our eyes provide us with an abundance of information about the outside world. Thanks to vision we become aware of the objects and living beings that surround us and represent their form and properties in our brains. Computer vision researchers aim at reproducing this capability in machines.[2]

Object Recognition

Object recognition is a computer vision technique which is used for identifying objects in images or videos. Object recognition is one of the most important outputs of machine learning and deep learning.

Object detection and object recognition are similar techniques but they vary in their execution. Object Detection algorithms act as a combination of both image classification

as well as object localization. This allows for multiple objects to be identified and located within the same frame.

The following section explains various approaches of machine learning and deep learning for object recognition.

Machine learning techniques are popular for object recognition and provide various different approaches than deep learning. Common examples of machine learning techniques are:

- **HOG (Histogram of oriented Gradients) feature Extractor and SVM (Support Vector Machine) model**: It takes histogram descriptors of both positive (those images which contain object) and negative (that image that does not contain objects) samples and trains our SVM model on that.
- **Bag of features model:** This approach basically represents an image as a collection of image features. Examples of this are SIFT, MSER, etc.
- **Viola-Jones algorithm:** This algorithm is mostly used for face detection in the image or real-time projects.

Deep learning techniques are the most popular methods for doing object recognition. Deep learning has three basic approaches: The Standard Feed forward neural network, RNN/LSTM, and Convolutional NN (CNN).

Convolution Neural Network (CNN) is the best approach from all the basic 3 approaches. The explanation for CNN is done with the help of black and white images as each pixel has only one value (from 0-255). The network layer for CNN is different from other neural networks. The four layers are: the convolution, the ReLUs, the pooling, and the fully connected layers.

1) *Convolution*:

The first step is to create many small pieces which are called features. Each feature resembles some part of the original image and then the features search through the original image. Once there is a high score there's a high score in that box, and if there's no match or low match then the score is zero or low. This process in producing the scores is referred to filtering

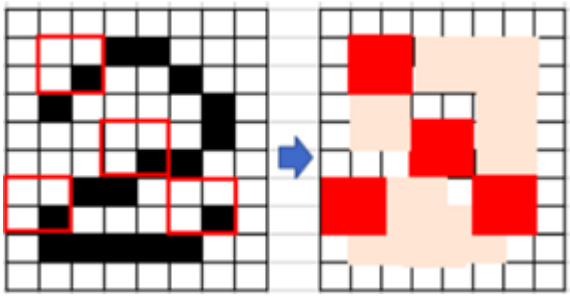


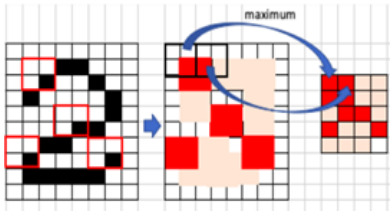
Image the feature has found similar matches in the original image. The red filled boxes are the high score and the pink boxes are low scores. The act of trying every possible match by scanning through the original image is called convolution.[2] All the filtered images are collected together and are called the convolution layer.

2) ReLUs layer:

The only task of this layer is to rectify any negative value to zero, so the maths behind the algorithm will work properly.

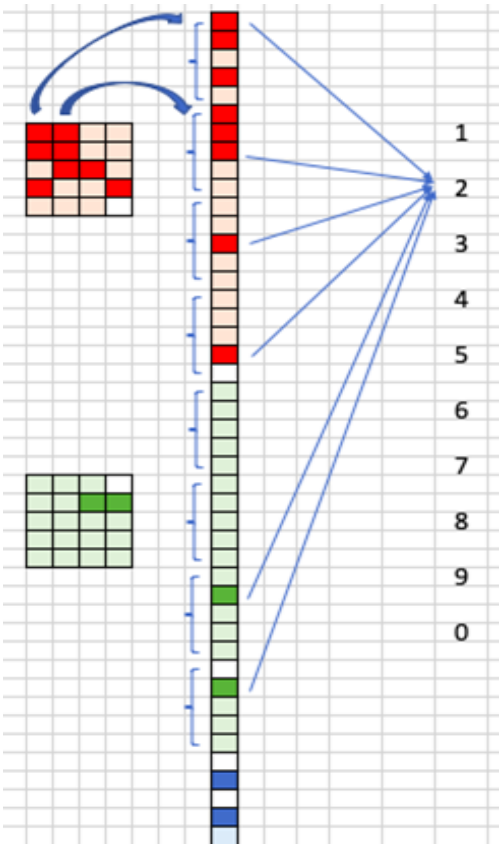
3) Max Pooling layer:

This process shrinks the image size. This collects the feature and picks the one unit with the highest score and assigns it to a new stack. The new stack of smaller featured images is produced



4) Fully connected layer (The final layer):

In this process we split the smaller featured image and stack them into a single list. This is the same process as in other typical neural networks. If the input is 2 and the 2 receives the highest score among all the nodes then the system recognizes the input as 2.



4. RESULTS

In a nutshell, keep in mind while choosing between machine learning and deep learning whether you have a powerful GPU and thousands of labeled training images. Deep learning techniques often perform better with more images, and a powerful GPU to decrease the time needed to train the model.

Acknowledgements

The completion of this review paper could not have been possible without the continued & dedicated efforts & guidance of Mrs. Gurpreet Kaur, faculty members of the Institute. We acknowledge our gratitude to her.

REFERENCES

1. https://en.wikipedia.org/wiki/Computer_vision#History
2. https://towardsdatascience.com/module-6-image-recognition-for-insurance-claim-handling-part-i-a338d16c9de0
3. https://in.mathworks.com/solutions/image-video-processing/object-recognition.htm

