IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

Computer Vision: The Era of Automation

Srinjoy Saha[1], Shreya Bhar[2], Sukalyan Porey[3], Kusum Majhi[4], Rupsa Das[5], Debrupa Pal[6]
Students, Department of Computer Application[1,2,3,4,5]
Assistant Professor, Department of Computer Application[6]
Narula Institute of Technology, Kolkata, India

Abstract: *In spite of its fancy sounding name, cloud computing is rather straightforward. The majority of data that we manage and store in our everyday lives is stored and managed via a few computing platforms. In the technological era, IT customers can access computing resources via the cloud, which is a service model used by information technology (IT). The cloud is increasingly becoming the preferred method of storing business data. This paper is mainly focuses on essential aspects of Computer Vision, like what is Computer Vision, Algorithms of Computer Vision, Applications, Challenges and Future Scope of Computer Vision.*

Keywords: Computer Vision, Artificial Intelligence (AI), 3D Modelling, Automation, Haar-like Features.

I. INTRODUCTION

With the advancement of artificial intelligence and machine learning, it became possible to develop machine systems with human-like intelligence in the early days. An artificial intelligence form called computer vision simulates and interprets the way humans understand and perceive the world around them. Multidimensional data can be collected via computer vision projects based on digital visual content. A computer interprets digital images and videos based on machine learning models, then identifies and classes objects accordingly. Artificial intelligence that teaches machines how to collect information from pixels is known as this branch of AI. Using this data, decision-makers can make use of it in computer-readable form.By combining virtual objects with real images, augmented-reality applications enable self-driving cars to navigate streets and highways safely, facial recognition tools identify people by matching faces to their identities and autonomous vehicles can navigate along streets safely. Defective products are not shipped to customers after they have been spotted on the assembly line [1][2].

II. ALGORITHMS

SURF Algorithm: It is a method for detecting and describing images with the help of Speeded Up Robust Features (SURF). As a result of its low dimensionality and lower computation time, SURF is faster than SIFT in real-time computer vision applications. For identifying objects and registering images, it is a robust and fast algorithm.

In situations where there is an illumination problem in the images, Surf is not stable to rotation. Instead of averaging the images using Gaussian distributions, SURF approximates the DoG with box filters. Using the integral image for convolution gives better results since it is faster [3].

There are two steps in this algorithm, those are -

1) Feature Extraction
2) Feature Description

SIFT Algorithm: A digital image's local features are detected and described using SIFT or the scale-invariant feature transform algorithm. The University of British Columbia has patented its use, however, which limits its usage. A descriptor is used for detecting and recognizing objects by locating key points and providing quantitative information about them.

In this algorithm, due to their local characteristics, these features are resistant to occlusions and clutter. The primary disadvantage of SIFT is its computationally intensive nature due to its high dimensionality. It performs poorly when the illumination changes, and is slow when the illumination changes [4].

There are four steps in this algorithm, those are -

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

1) Scale-space Extrema Detection
2) Key point Localization
3) Orientation Assignment
4) Keypoint Descriptor

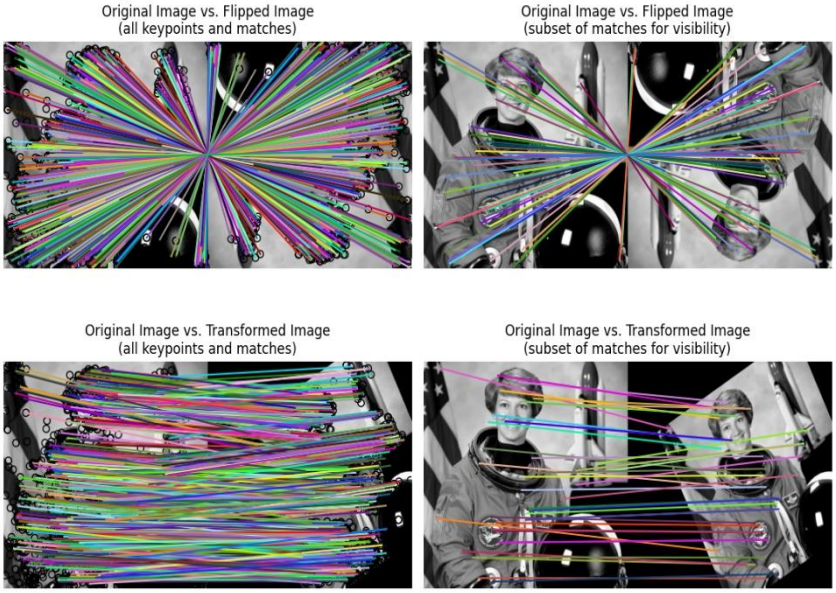

Figure 1: SIFT Algorithm

Viola-Jones Algorithm: Researchers Paul Viola and Michael Jones developed the Viola-Jones object detection algorithm in 2001 in order to solve the problem of face detection. The algorithm uses a technique called "integral image" to calculate Haar features quickly. In applications such as security systems, photo tagging and other ones that do not require a great deal of computational power, the Viola-Jones algorithm is particularly efficient.

As training datasets get larger, the algorithm can become more difficult to train, but real-time face detection remains quite fast [5][6]. Although Viola-Jones is one of the longest running face detection algorithms, it produces outstanding results when applied to real-time face detection.

In this algorithm there are four main steps, those are -

1) Selecting Haar-like Features
2) Creating an Integral Image
3) Running AdaBoost Training
4) Creating Classifier Cascades

Lucas-Kanade Algorithm: In optical flow estimation, one of the most widely used algorithms is the Lucas-Kanade algorithm, which identifies pixel-level movements between consecutive photographs. According to the algorithm, pixels in the local area of an image have a constant optical flow. Based on the assumption that brightness remains constant, Lucas Kanade developed his algorithm.

Lucas Kanade method further assumes that pixel values will move similarly inside an object as they move in a scene. Other object tracking algorithms are more difficult to implement than Lucas Kanade. Calculating and predicting the motion of objects can be done with good accuracy and calculation speed. Despite being able to handle moderate object speeds, the Lucas Kanade algorithm has the disadvantage of not performing well in rapid motion [7].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

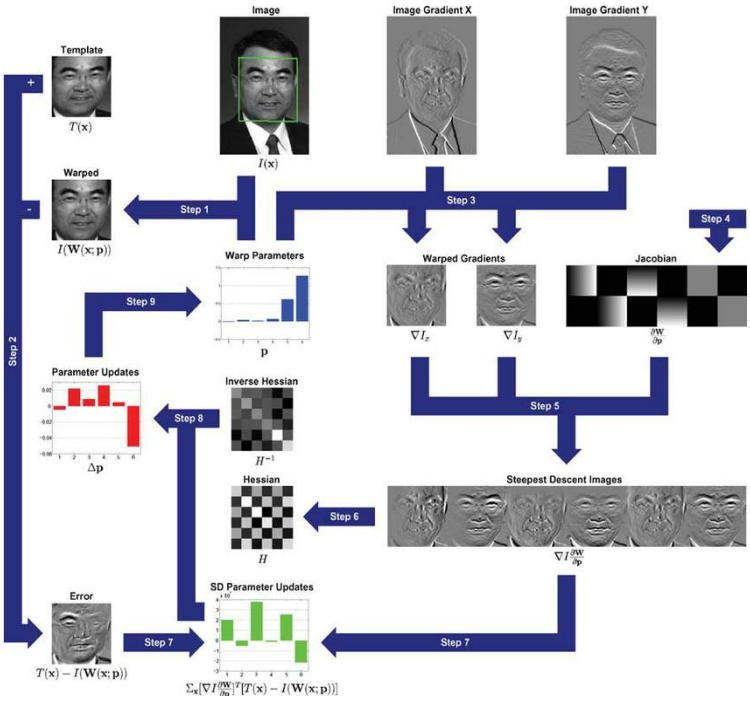

Figure 2: Lucas-Kanade Algorithm

III. APPLICATIONS

- **Face Recognition:** A major application of computer vision in facial recognition programs is recognizing individuals in photographs using computer vision. A computer vision algorithm detects facial features on images and compares them with fake profiles based on input data, which is initially presented to the machines as input data. Detecting and tagging users is both possible with facial recognition in social networking applications like Facebook, Instagram. This feature is also being used by government spy agencies to identify criminals in videos.
- **Healthcare Sector:** Technology in the health care field has been greatly influenced by computer vision. Computer vision technology provides fast and accurate assessments of chemotherapy response to cancerous tumours than traditional approaches, which take time and have poor accuracy predictions. By identifying cancer patients who need surgery more quickly, doctors can save lives. In addition to finding indicators in x-rays or MRI scans, computer vision algorithms can also be used to automatically recognize malignant moles on the skin. The use of computer vision, especially in the area of organ and tissue visualization, helps medical professionals make better decisions regarding patient treatment [8][9].
- **Self-Driving Cars:** Computer vision is a crucial AI field for self-driving cars. It allows computers to understand visual info from sensors (cameras, lidar, radar). Semantic segmentation, depth estimation, tracking and motion estimation, and lane detection are some of the computer vision techniques used in self-driving cars. Computer vision is also used for traffic sign recognition, pedestrian detection, and behaviour prediction. Large-scale datasets are essential for training and validating computer vision algorithms for self-driving cars. Hardware captures data and software processes it in real-time.
- **3D Modelling and Automotive Safety:** We can also use computer vision for developing 3D Models. The process of building a 3D model using software is known as 3D modelling. A 3D computer model can be constructed from existing objects using computer vision in this field as well. Aside from robotics and autonomous driving, 3D modelling is used in 3D scene reconstruction, 3D tracking, and augmented reality. On the other hand, a key safety feature of the automotive industry has been introduced by computer vision. It

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

might prevent thousands of deaths and property damage if a vehicle is taught to detect obstacles and dangers [10] [11].

IV. CHALLENGES

Computer vision is one of the most emerging and growing domains of AI, but still there are few challenges which can be major hindrance in the way of becoming a leading technology. These challenges are:

- **Reasoning and Analytical Issues:** All types of programming languages and technologies require the basic logic behind any task. To become a computer vision expert, one must have good reasoning and analytical skills, otherwise defining any attribute in visual content may be a big problem.
- **Real-Time Processing:** There are many computer vision applications like autonomous driving, robotics etc. require real-time processing of visual data. To achieve high-speed data processing and maintaining low latency is crucial in these applications. It poses a big challenge due to the complexity of the algorithms and the estimated requirements [12].
- **Privacy and Security:** Privacy and security are one of the most important factors. Computer Vision surveillance is also having various kind of privacy issues that restricts users from accessing the unauthorized content. So, many countries avoid such face recognition and detection techniques for privacy and security problems.
- **Duplicate and False Content:** Duplicate and false contents pose a big challenge for computer vision systems. Duplicate content refers to identical or highly similar images or videos which can be used to deceive computer vision systems. On the other hand, False content can include computer-generated images or videos that credibly mimic real-world people or events. Detecting and distinguishing between real and fake content can be a serious challenge for computer vision systems [13] [14].

V. FUTURE SCOPE & CONCLUSION

Computer vision will perform more tasks in the future. It will be used for visually impaired people and to develop AGI and ASI. Computer vision AI at the edge is essential for real-time responsiveness. It can improve performance and enable real-time continuous improvement. Computer vision has more benefits and applications to explore. The biggest hurdle is demystifying the black box of AI. Validate computer vision models and monitor them for security, reliability, privacy, and ethics. Ethics is a starting point for frameworks.

Computer vision is an emerging field with significant potential that can transform various industries and applications. It is a field of computer science that uses tools like deep learning and convolutional neural networks to develop various methods to help computer systems to see and understand different types of media, such as photos and videos. Now computers systems can recognize thousands of new faces as well as objects with less time and more accuracy [15].

One of the major advantages of computer vision is the emergence of deep learning and convolutional neural networks. Before the emergence of deep learning, computer vision systems could only perform limited tasks that required a lot of coding and effort by developers. Deep learning models have made remarkable success in tasks like image recognizing, object detection, in some cases it has surpasses performances of humans.

While computer vision has given remarkable performance in recent years, but still, it faces several challenges. Some featured areas of ongoing research include handling variations in lighting conditions, blockage, and changes in viewpoint, improvement of interpretability and explain ability of deep learning models, improving privacy and security [16].

REFERENCES

[1] Y. Yakimovsky and R. Cunningham, “A system for extracting three-dimensional measurements from a stereo pair of TV cameras," Compute. Graphics Image Processing, vol. 7, pp. 195-210, 1978.

[2] H. P. Moravec, "Visual mapping by a robot rover," in Proc. 6th Int. Joint Conf. Artificial Intell., 1979, pp. 598-620.

[3] H. H. Baker, "Edge based stereo correlation," in Proc. ARPA Image Understanding Workshop, Univ. Maryland, Apr. 1980.

[4] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," M.I.T., A.I. Lab., Memo. 364, June 1976.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, July 2023

[5] B. Julesz, "Binocular depth perception without familiarity cues," Science, vol. 145, pp. 356-362, 1964.
[6] D. Marr and T. Poggio, "Computational approaches to image understanding," M.I.T., A.I. Lab., see also Proc. R. Soc. London B, vol. 204, pp. 301-328, 1979.
[7] F. Röcker and A. Kiessling, "Methods for analysing three dimensional scenes," in Proc. 4th Int. Joint Conf. Artificial Intell., 1975, pp. 669-673.
[8] B. K. P. Horn, "Focussing," M.I.T., Project MAC, AI Memo. 160, May 1968.
[9] R. A. Jarvis, "Focus optimisation criteria for computer image processing," Microscope, vol. 24, pp. 163-180, 2nd quarter, 1976.
[10] B. K. P. Horn, "Shape from shading: A method for obtaining the shape of a smooth opaque object from one view," M.I.T., Project MAC, MAC TR-79, Nov. 1970.
[11] K. Ikeuchi and B. K. P. Horn, "An application of the photometric stereo method," in Proc. 6th Int. Joint Conf. Artificial Intell., Tokyo, Japan, 1979, pp. 413-415.
[12] M. L. Braunstein, Depth Perception Through Motion. New York: Academic, 1976.
[13] Y. Shirai and M. Suwa, "Recognition of polyhedrons with a range finder," in Proc. 2nd Int. Joint Conf. Artificial Intell,, London, Sept. 1971, pp. 80-87.
[14] M. D. Levine, D. A. O'Handley, and G. M. Yagi, "Computer determination of depth maps," Compute. Graphics Image Processing, vol. 2, pp. 134-150, 1973.
[15] R. A. Jarvis, "Optimisation in adaptive control: A selective survey," IEEE Trans. Syst., Man, Cybern., vol. SMC-5, pp. 83-94, Jan. 1975.
[16] G. J. Agin and T. O. Binford, "Computer description of curved objects," in Proc. Int. Joint Conf. Artificial Intell., Stanford Univ., Aug. 20-23, 1973, pp. 629-640.