

COMPUter VIsion



Artificial Intelligence Department Faculty of Computer science and artificial intelligence

Computer Vision

Technical report for technical writing course in Faculty of Computer science and artificial intelligence code 5.

Abstract.

In this report, the concept of computer vision will be explained, along with examples such as automatic cars, and the purpose of computer vision, such as why we should care about this field, for instance, to teach machines how to gather information from pixel units. It will explain how computer vision works, relying on self-training pattern recognition. The core functions of computer vision, such as OCR, mocap, medical imaging, and others, will be elucidated in the article. The beneficial outcomes of learning computer vision will be highlighted, such as becoming a computer vision engineer, which will have significant demand in the current and future market. Additionally, the languages used and most suitable for computer vision, notably Python, will be clarified. The article will also explain the practical results of computer vision applications, such as their presence in smartphones and other aspects of our daily lives. Ultimately, it will be noted that the future of computer vision will spread to a wider range of content.

Acknowledgement.

I would like to thank Dr /Noha Yehia for her assistance in preparing and organizing this report and her efforts to foster a spirit of cooperation among the team as both students now and future colleagues, as well as her efforts to facilitate communication among us and develop communication skills.

I would also to express my sincere gratitude to the team members for their valuable contributions and support throughout the preparation of this report. Their expertise, encouragement, and assistance were instrumental in the collaboration among us, helping each other to complete this report, and their mutual respect and appreciation.

Their collective efforts significantly contributed to the quality and depth of this report.

Table of Contents

**Abstract**………………………………………………………………………………………………i

Acknowledgements…………………………………………………………………………………...ii

List of figures…………………………………………………………………………………………iv

**Introduction**…………………………………………………………………………………………..**1**

Central topic:

Origin of computer vision………………………………………………………………......2

Why computer vision is important……………………………………………………………………..2

How does computer vision work…………………………………………………………….3

Computer vision advantages and disadvantages………………………………………………....4

Computer Vision Application………………………………………………………………..5

How to learn Computer Vision?.............................................................................................6

Which language is best suited for computer vision?...............................................................7

Challenges we face in Computer Vision…………………………………………………………….8

Future of Computer Vision……………………………………………………………………………...8

Conclusion………………………………………………………………………………………………………………9

References………………………………………………………………………………………………………………10

List of figures

[Figure 1computer vision. 4](#_Toc165892210)

Figure 2 *The difference between the way human vision and computer vision work………………….*

Figure 3 computer vision benefits……………………………………………………………………………………..

Figure 4 lack of monitoring is one of disadvantages……………………………………………………………...

Figure 5 self-driving car application……………………………………………………………………………………

Figure 6 facial recognise.

application…………………………………………………………………………………

Figure 7 Augmented & Mixed Reality application…………………………………………………………….

Figure 8 Health care application……………………………………………………………………………………

Figure 9 one of ways to learn computer vision………………………………………………

Figure 10 python is the best language to learn computer vision………………………….

Figure 11 one of challenges which face computer vision……………………………………

Introduction

if you were told to name certain things that you’d find in a park, you’d casually mention things like grass, bench, trees, etc. This is a very easy task that any person can accomplish in the blink of an eye. However, there is a very complicated process that takes place in the back of our minds.

Human vision involves our eyes, but it also involves all our abstract understanding of concepts and personal experiences through millions of interactions we have had with the outside world. Until recently, computers had very limited abilities to think independently. Computer vision is a recent branch of technology that focuses on replicating this human vision to help computers identify and process things the same way humans do.

Computer vision is a field of study which enables computers to replicate the human visual system. It’s a subset of artificial intelligence which collects information from digital images or videos and processes them to define the attributes. The entire process involves image acquiring, screening, analysing, identifying and extracting information. This extensive processing helps computers to understand any visual content and act on it accordingly. You can also take up a computer vision course for free to understand the basics under Artificial intelligence domain.  
Computer vision projects translate digital visual content into explicit descriptions to gather multi-dimensional data. This data is then turned into computer-readable language to aid the decision-making process. The main objective of this branch of artificial intelligence is to teach machines to collect information from pixels. The field of computer vision has made significant progress. Because this field is important, we will learn about it in some detail.

Central topic

What is Computer Vision?

Computer vision is one of the fields of artificial intelligence that trains and enables computers to understand the visual world. Computers can use digital images and deep learning models to accurately identify and classify objects and react to them.

Computer vision in AI is dedicated to the development of automated systems that can interpret visual data (such as photographs or motion pictures) in the same manner as people do. The idea behind computer vision is to instruct computers to interpret and comprehend images on a pixel-by-pixel basis. This is the foundation of the computer vision field. Regarding the technical side of things, computers will seek to extract visual data, manage it, and analyze the outcomes using sophisticated software programs.

The amount of data that we generate today is tremendous - 2.5 quintillion bytes of data every single day. This growth in data has proven to be one of the driving factors behind the growth of computer vision.

History

For almost 60 years, researchers and developers have sought to teach computers how to perceive and make sense of visual information. In 1959, neurophysiologists started showing a cat a variety of sights to correlate a reaction in the animal's brain. They found that it was particularly sensitive to sharp corners and lines, which technically indicates that straight lines and other basic forms are the foundation upon which image analysis is built.

Around the same period, the first image-scanning technology emerged that enabled computers to scan images and obtain digital copies of them. This gave computers the ability to digitize and store images. In the 1960s, artificial intelligence (AI) emerged as an area of research, and the effort to address AI's inability to mimic human vision began.

Neuroscientists demonstrated in 1982 that vision operates hierarchically and presented techniques enabling computers to recognize edges, vertices, arcs, and other fundamental structures. At the same time, data scientists created a pattern-recognition network of cells. By the year 2000, researchers were concentrating their efforts on object identification, and by the following year, the industry saw the first-ever real-time face recognition solutions.

Origin of Computer Vision:

Computer vision is not a new concept; in fact, it dates to the 1960s. It all started with an MIT project - “Summer Vision Project” which analysed scenes to identify objects. David Marr, the celebrated neuroscientist, laid down the building blocks of computer vision, taking a cue from the functions of the cerebellum, hippocampus, and cortex of human perception. He has been dubbed the father of computer vision since, and the field has evolved to include much more complicated functionalities.



Figure computer vision.

Why is Computer Vision Important?

From selfies to landscape images, we are flooded with all kinds of photos today. According to a report by Internet Trends, people upload more than 1.8 billion images every day, and that’s just the number of uploaded images. Imagine what the number will come to if you consider the images stored in phones. We consume more than 4,146,600 videos on YouTube and send 103,447,520 spam mails every day. Again, that’s just a part of it – communication, media and entertainment, the internet of things are all actively contributing to this number. This abundantly available visual content demands analysing and understanding. Computer vision helps in doing that by teaching machines to “see” these images and videos.

Additionally, thanks to easy connectivity, the internet is easily accessible by all today. Children are especially susceptible to online abuse and “toxicity”. Apart from automating a lot of functions, computer vision also ensures moderation and monitoring of online visual content. One of the main tasks involved in online content curation is indexing. Since the content available on the internet is mainly of two types, namely text, visual, and audio categorisation becomes easy. Computer vision uses algorithms to read and index images. Popular search engines like Google and YouTube use computer vision to scan through images and videos to approve them for featuring. By way of doing so, they not only provide users with relevant content but also protect against online abuse and “toxicity”.

How Does Computer Vision Work?

Massive amounts of information are required for computer vision. Repeated data analyses are performed until the system can differentiate between objects and identify visuals. Deep learning, a specific kind of machine learning, and convolutional neural networks, an important form of a neural network, are the two key techniques that are used to achieve this goal.

With the help of pre-programmed algorithmic frameworks, a machine learning system may automatically learn about the interpretation of visual data. The model can learn to distinguish between similar pictures if it is given a large enough dataset. Algorithms make it possible for the system to learn on its own, so that it may replace human labour in tasks like image recognition.

Convolutional neural networks aid machine learning and deep learning models in understanding by dividing visuals into smaller sections that may be tagged. With the help of the tags, it performs convolutions and then leverages the tertiary function to make recommendations about the scene it is observing. With each cycle, the neural network performs convolutions and evaluates the veracity of its recommendations. And that's when it starts perceiving and identifying pictures like a human.

Computer vision is like solving a jigsaw puzzle in the real world. Imagine that you have all these jigsaw pieces together and you need to assemble them to form a real image. That is exactly how the neural networks inside a computer vision work. Through a series of filtering and actions, computers can put all the parts of the image together and then think on their own. However, the computer is not just given a puzzle of an image - rather, it is often fed with thousands of images that train it to recognize certain objects.

For example, instead of training a computer to look for pointy ears, long tails, paws and whiskers that make up a cat, software programmers upload and feed millions of images of cats to the computer. This enables the computer to understand the different features that make up a cat and recognize it instantly.

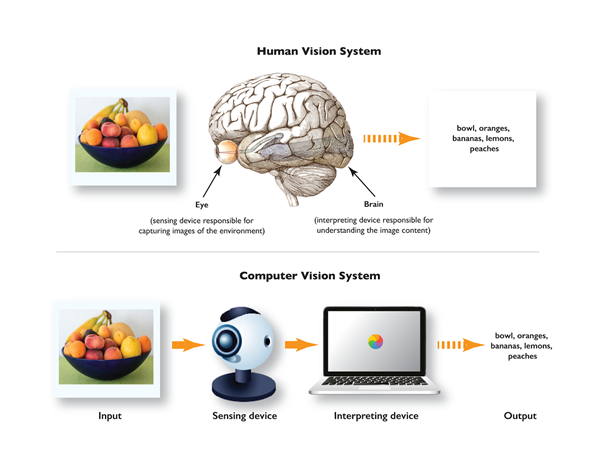


Figure2: The difference between the way human vision and computer vision work

Computer Vision Advantages

Computer vision can automate several tasks without the need for human intervention. As a result, it provides organizations with several benefits:

Faster and simpler process - Computer vision systems can carry out repetitive and monotonous tasks at a faster rate, which simplifies the work for humans.

Better products and services - Computer vision systems that have been trained very well will commit zero mistakes. This will result in faster delivery of high-quality products and services.

Cost-reduction - Companies do not have to spend money on fixing their flawed processes because computer vision will leave no room for faulty products and services.

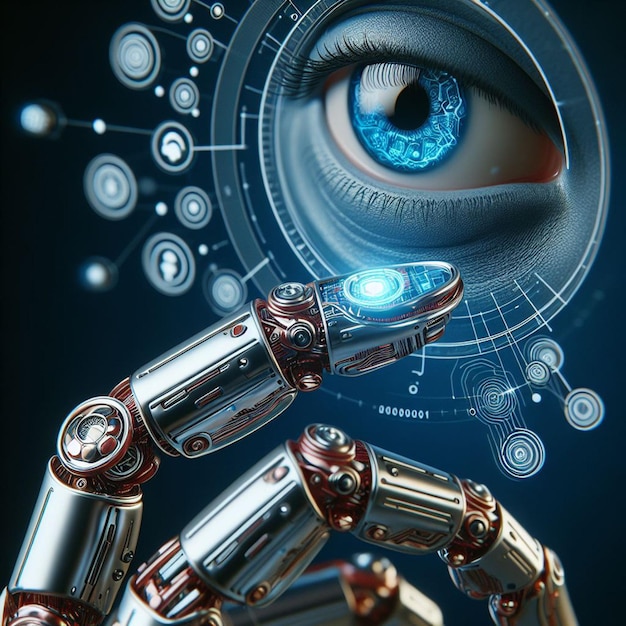


Figure3 computer vision benifits

Computer Vision Disadvantages

There is no technology that is free from flaws, which is true for computer vision systems. Here are a few limitations of computer vision:

Lack of specialists - Companies need to have a team of highly trained professionals with deep knowledge of the differences between AI vs. Machine Learning vs. Deep Learning technologies to train computer vision systems. There is a need for more specialists that can help shape this future of technology.

Need for regular monitoring - If a computer vision system faces a technical glitch or breaks down, this can cause immense loss to companies. Hence, companies need to have a dedicated team on board to monitor and evaluate these systems.



Figure 4 lack of monitoring is one of disadvantages.

Computer Vision Applications

One field of Machine Learning where fundamental ideas are already included in mainstream products is computer vision. The applications include:

1.Self-Driving Cars

With the use of computer vision, autonomous vehicles can understand their environment. Multiple cameras record the environment surrounding the vehicle, which is then sent into computer vision algorithms that analyses the photos in perfect sync to locate road edges, decipher signposts, and see other vehicles, obstacles, and people. Then, the autonomous vehicle can navigate streets and highways on its own, swerve around obstructions, and get its passengers where they need to go safely. 

Figure 5 self-driving car application

2.Facial Recognition

Facial recognition programs, which use computer vision to recognize individuals in photographs, rely heavily on this field of study. Facial traits in photos are identified by computer vision algorithms, which then match those aspects to stored face profiles. To verify the identity of the people using consumer electronics, face recognition is increasingly being used. Facial recognition is used in social networking applications for both user detection and user tagging. For the same reason, law enforcement uses face recognition software to track down criminals using surveillance footage.



Figure 6 facial recognise application.

3.Augmented & Mixed Reality

Augmented reality, which allows computers like smartphones and wearable technology to superimpose or embed digital content onto real-world environments, also relies heavily on computer vision. Virtual items may be placed in the actual environment through computer vision in augmented reality equipment. To properly generate depth and proportions and position virtual items in the real environment, augmented reality apps rely on computer vision techniques to recognize surfaces like tabletops, ceilings, and floors.



Figure 7 Augmented & Mixed Reality application

4.Healthcare

Computer vision has contributed significantly to the development of health tech. Automating the process of looking for malignant moles on a person's skin or locating indicators in an x-ray or MRI scan is only one of the many applications of computer vision algorithms.



Figure 8 Healthcare application

How to learn Computer Vision?

Depending on the uses, computer vision has the following uses:

1.Laying the Foundation: Probability, statistics, linear algebra, calculus and basic statistical knowledge are prerequisites of getting into the domain. Similarly, knowledge of programming languages like Python and MATLAB will help you grasp the concepts better.

2.Digital Image Processing: Learn how to compress image and videos using JPEG and MPEG files. Knowledge of basic image processing tools like histogram equalisation, median filtering and more are required. Once you know the basics of image processing and restoration, you will be ready to pick up the more critical skills of computer vision.

3.Machine Learning Basics: Knowledge of Convoluted Neural Networks, fully connected neural networks, support vector machines, recurrent neural networks, generative adversarial network, and autoencoders are necessary to get started with computer vision.

4.Basic Computer Vision: The next step in the process is to decode the mathematical models involved in the image and video formulations. Once you understand how pattern recognition and signal processing works, you can get into advanced learning.



Figure 9 one of ways to learn computer vision.

Which language is best suited for computer vision?

Computer vision experts recommend Python for the following reasons:

Easy to Use: Python is easy to learn, especially for beginners. It is one of the first programming languages learnt by most users. This language is also easily adaptable for all kinds of programming needs.

Most Used computing language: Python offers a complete learning environment for people who want to use it for various kinds of Computer Vision and Machine Learning experiments. Its NumPy, scikit-learn, matplotlib and OpenCV provides an exhaustive resource for any computer vision applications.

Debugging and Visualisation: Python has an in-built debugger, ‘PDB’ which makes debugging codes in this programming language more accessible. Similarly, Matplotlib is a convenient resource for visualisation.

Web Backend Development: Frameworks like Django, Flask, and Web2py are excellent web page builders. Python is compatible with these frameworks and can be easily tweaked to fit your requirements.

MATLAB is the other programming language popular with computer experts. Let’s investigate the advantages of using MATLAB:

Toolboxes: MATLAB has one the most exhaustive toolboxes; whether it is a statistical and machine learning toolbox, or an image processing toolbox, MATLAB has one included for all kinds of needs. The clean interfaces of each of these toolboxes enables you to implement a range of algorithms. MATLAB also has an optimisation toolbox which ensures that all algorithms perform at their best.

Powerful Matrix Library: Images and other visual content contain multi-dimensional matrices along with linear algebra in different algorithms which becomes easier to work within MATLAB. The linear algebra routines included in MATLAB work fast and effective.

Debugging and Visualisation: Since there is a single integrated platform for coding in MATLAB, writing, visualising and debugging codes become easy.

Excellent Documentation: MATLAB enables you to document your work adequately so that it is accessible later. Documentation is essential not just for future reference but also to help coders work faster. MATLAB’s documentation allows users to work twice the speed of OpenCV.

Computer Vision experts also gravitate towards OpenCV for the follow.



Figure 10 python is the best language to learn computer vision.

**How to become a Computer Vision Engineer?**

Computer vision engineers are in high demand in the market today, thanks to the enormous amount of visual content that needs to be worked upon.

**What exactly does a Computer Engineer do?**

* A computer vision engineer creates and uses vision algorithms to work on the pixels of any visual content (images, videos and more)
* They use a data-based approach to develop solutions.
* They usually come with a background in AIML and have experience working on a variety of systems, including segmentation, machine learning, and image processing.  
  If you want to become a computer vision engineer, you need to pick up the basic skills of the domain and work on projects that will give you a hands-on experience of industry-relevant problem-solving. [Great Learning](https://www.mygreatlearning.com/)’s [Deep Learning certificate program](https://www.mygreatlearning.com/pg-program-artificial-intelligence-course) introduces you to all the basics of the domain and sets you on the path of becoming a computer vision engineer.

**Job Description of Computer Vision Engineer**

The ideal candidate must have a sound knowledge of machine learning algorithms, principles, and their application. He/she should have experience working on [Deep Learning](https://www.mygreatlearning.com/blog/what-is-deep-learning/) architectures like CNN, GAN, , and more. He/she should also be familiar with deep learning frameworks like TensorFlow and Porch. He/she must also have a good understanding of object detection and models like YOLO, RCNN, Mask-RCNN and more.

**Requirements in Computer Vision Engineers**

* Knowledge of process automation and AI pipeline designing.
* 1+ years of experience in Artificial Intelligence projects
* Programming skills (Python, C++, MATLAB) is a must.
* Ability to drive projects independently and with the team.
* Working knowledge of tools like git, docker etc.
* Excellent written and verbal communication skills
* Degrees in computer science, electrical engineering preferred

Figure 11 **Requirements in Computer Vision Engineers:**

Challenges we face in Computer Vision

Reasoning Issue: Modern neural network-based algorithms are complex system whose functioning’s are often obscure. In situations like these, it becomes tough to find the logic behind any task. This lack of reasoning creates a real challenge for computer vision experts who try to define any attribute in an image or video.

Privacy and Ethics: Vision powered surveillance is a serious threat to privacy in a lot of countries. It exposes people to unauthorised use of data. Face recognition and detection is prohibited in some countries because of these problems.

Fake Content: Like all other technologies, computer vision in the wrong hands can lead to dangerous problems. Anybody with access to powerful data centres can create fake images, videos or text content.

Adversarial Attacks: These are optical illusions for the computer. When an attacker creates a faulty machine learning model, they intend the machine using it to fail. These flawed models are difficult to identify and can cause serious damage to any system.



Figure 11 one of challenges which face computer vision.

Future of Computer Vision

Computer vision is a fast-developing field and has gathered a lot of attention from various industries. It will be able to function on a broader spectrum of content in the future. The domain already enjoys a steady market of 2.37 million US dollars and is expected to grow at a 47% CAGR till 2023. With the amount of data we are generating every day, it’s only natural that machines will use that data to craft solutions.

Once computer vision experts can resolve the current problems of the domain, we can expect a trustworthy system that automates content moderation and monitoring. With corporate giants like Google, Facebook, Apple and Microsoft investing in computer vision, it’s only a matter of time before it takes over the global market. Upskill in this domain to make the most of this disruptive economy. Thus, we wrap up with this quick introduction to computer vision. Hope you enjoyed the blog and if you did, please share and comment your thoughts below.

Conclusion

After we learned about the field of computer vision, its importance, and how to learn it, and we learned about the challenges facing this field and realized its importance, we must learn this field through Laying the Foundation: Probability, statistics, linear algebra, calculus and basic statistical knowledge are prerequisites of getting into the domain. Similarly, knowledge of programming languages like Python and MATLAB will help you grasp the concepts better, Digital Image Processing: Learn how to compress image and videos using JPEG and MPEG files. Knowledge of basic image processing tools like histogram equalisation, median filtering and more are required. Once you know the basics of image processing and restoration, you will be ready to pick up the more critical skills of computer vision, Machine Learning Basics: Knowledge of Convoluted Neural Networks, fully connected neural networks, support vector machines, recurrent neural networks, generative adversarial network, and autoencoders are necessary to get started with computer vision, And Basic Computer Vision: The next step in the process is to decode the mathematical models involved in the image and video formulations. Once you understand how pattern recognition and signal processing works, you can get into advanced learning, and we well use python because it is the best language for learning computer vision. Despite its successes, computer vision faces several challenges. Robustness and generalization are key concerns, as models trained on specific datasets may struggle with real-world variability. Additionally, ethical and privacy issues, especially in areas like surveillance and biometric identification, require thoughtful regulation and safeguards to prevent misuse or overreach. Looking ahead, computer vision is poised for continued growth, with expanding use cases in augmented reality, virtual reality, and smart cities. Innovations in hardware, like specialized AI chips and edge computing, are likely to enhance the performance and scalability of computer vision systems. However, the ethical implications of these technologies must be carefully managed, emphasizing transparency, accountability, and respect for privacy.

In summary, computer vision is a dynamic and rapidly evolving field that has made significant strides in recent years. Its potential to revolutionize industries and impact daily life is substantial, but ongoing efforts are needed to address the technical and ethical challenges it presents. By focusing on responsible development and deployment, computer vision can continue to offer transformative benefits while minimizing risks.

References

# Reference

simplilearn. (2024, may 7). the power of computer vision in Ai. *computer vision* , pp. 20-35.

# Reference

Chatterjee, M. (2024, may 6). *What is Computer Vision? Know Computer Vision Basic to Advanced & How Does it Work?* Retrieved from Great learning: https://www.mygreatlearning.com/blog/what-is-computer-vision-the-basics/

# Reference

altman, s. (n.d.). *acknowledgements* . Retrieved from chat gbt: http://chatgbt.com

# Reference

altman, s. (n.d.). *abstract*. Retrieved from chat gbt: http://chatgbt.com.