Computer Vision and AI

\*Note: Sub-titles are not captured in Xplore and should not be used

Alihan Esen   
AI and Data Engineering  
Istanbul Technical UniversityIstanbul, Turkey  
esenal22@itu.edu.tr

Ahmet Enes Topçu   
AI and Data Engineering  
Istanbul Technical UniversityIstanbul, Turkey  
topcua21@itu.edu.tr

Emirhan Gazi  
AI and Data Engineering  
Istanbul Technical UniversityIstanbul, Turkey  
gazi20@itu.edu.tr

line 1: 5th Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCIDTunahan Görgülü   
AI and Data EngineeringIstanbul Technical UniversityIstanbul, Turkey  
gorgulut21@itu.edu.tr

line 1: 6th Given Name Surname  
line 2: *dept. name of organization   
(of Affiliation)*  
line 3: *name of organization   
(of Affiliation)*line 4: City, Country  
line 5: email address or ORCID

*Abstract*—In this report we talk about how and where computer vision is used, what are the similarities of computer vision to human vision, who are the pioneers of computer vision and what were their influences in the field.

Keywords—Computer Vision, Artificial Intelligence, Cameras.

# What is Computer Vision?

Computer vision is a subfield of AI that focuses on enabling computers to understand and interpret visual data, such as images and videos. It allows computers to process and analyze this data in order to take actions or make recommendations based on what they have "seen." In this way, while AI as a whole enables computers to think and perform tasks independently, computer vision specifically allows computers to interpret and understand the visual world around them.

Computer vision and human vision share many similarities, with both being able to interpret and understand visual data. However, human vision has the benefit of being developed and refined over a lifetime of learning and experience, allowing us to differentiate between objects, gauge distances, detect movement, and identify abnormalities in images. Computer vision, on the other hand, must rely on cameras, data, and algorithms to achieve similar capabilities, although it can do so with a higher degree of precision and accuracy. While humans rely on our eyes, optic nerves, and visual cortex to process visual information, computer vision systems utilize specialized equipment and algorithms to accomplish these tasks in a shorter amount of time.

While humans may be skilled at identifying and interpreting visual information, we are limited in our ability to do so continuously and at a rapid pace. In contrast, computer vision systems that are designed to inspect products or monitor production assets can analyze thousands of items or processes in a minute, detecting even subtle defects or issues. In this way, computer vision systems can outperform human capabilities in terms of speed and continuous analysis.

# Pioneers of Computer Vision

Larry Roberts is widely recognized as the father of computer vision for his work on extracting 3D geometrical information from 2D perspective views of polyhedra in his Ph.D. thesis at MIT in the 1960s. This work sparked interest among researchers in the field of AI, both at MIT and elsewhere, who continued to study computer vision in the context of the "blocks world" and its applications.

As research in the field of computer vision progressed, it became apparent that it was necessary to address the challenges of interpreting images from the real world. This led to a focus on "low-level" vision tasks, including edge detection and segmentation. One significant advancement in this area was the framework proposed by David Marr at MIT in the 1970s, which adopted a bottom-up approach to understanding and interpreting scenes.

In the process of interpreting images, low-level image processing algorithms are often applied to 2D images to extract basic information, such as directed edge segments. This information is then used to create a 2.5D sketch of the scene using a binocular stereo. Finally, high-level techniques that incorporate structural analysis and prior knowledge are used to create 3D model representations of the objects in the scene. The framework proposed by David Marr is widely regarded as one of the most influential works in the field of computer vision.

# Traffic Drones

Traffic drones, also known as drone traffic management systems, are unmanned aerial vehicles (UAVs) that are used to monitor and control traffic on roads and highways. They are equipped with sensors and cameras that can detect accidents, congestions, and other traffic-related issues, and they can transmit this information back to traffic controllers in real-time. Some traffic drones are also equipped with loud speakers and LED lights, which can be used to direct traffic and alert drivers of potential hazards. Traffic drones are becoming increasingly popular as a way to improve traffic flow and reduce congestion, as well as to make roads safer for drivers, pedestrians, and cyclists.

## Ethical Considerations

There are several ethical considerations surrounding the use of traffic drones. One concern is privacy. Some people may be uncomfortable with the idea of drones flying overhead and collecting data on their movements and activities. To address this concern, it is important that traffic drones be used in a transparent and accountable manner, and that clear policies be put in place to ensure that the data they collect is not misused.

Another ethical consideration is the potential for drones to be used for nefarious purposes, such as surveillance or espionage. To prevent this, it is important that traffic drones be designed and deployed in a way that minimizes the risk of hacking or other forms of tampering.

In addition, there are concerns about the safety of traffic drones, both for the people who operate them and for those on the ground. To address these concerns, it is important that traffic drones be designed and built to high standards of safety and reliability, and that they be operated by trained and competent personnel.

Overall, the ethical use of traffic drones requires a careful balance between the benefits they can provide and the potential risks and concerns they may raise. It is important that any decisions about the use of traffic drones be made with careful consideration of these ethical issues.

# Retaıl Industry

The field of computer vision is fast expanding and has had a big impact on the retail sector. In order to extract useful information from photographs and videos, it uses computer algorithms to evaluate and comprehend the data. Computer vision has been applied in a variety of ways in the retail industry to enhance customer satisfaction and boost productivity.

Personalized recommendations and targeted advertising are two significant applications of computer vision in the retail sector. Retailers can utilize computer vision to provide customized product recommendations and targeted marketing by examining client behavior and preferences. Both the customer experience and sales can be enhanced by this.

In the retail industry, computer vision has also been applied to inventory control. Retailers may keep track of inventory levels and determine when products need to be refilled by examining photographs of products on store shelves. This can increase productivity and lessen the possibility of stock shortages.

Tracking consumer behavior is another retail industry use for computer vision. Retailers may learn how customers move throughout the store, which products they are interested in, and how much time they spend in various places by watching video of customers in stores. The layout of stores can be optimized using this data, which would also enhance the entire shopping experience.

# Mılıtary Industry

The field of computer vision is one that is expanding quickly and has had a big impact on the military. In order to extract useful information from photographs and videos, it uses computer algorithms to evaluate and comprehend the data. Computer vision has been utilized in the military for many different purposes, such as target recognition, surveillance, and tracking of moving objects like cars and people. Unmanned aerial vehicles have also used it to enhance situational awareness and aid in navigation.

Target recognition is a crucial use of computer vision in the military. Military personnel can swiftly and precisely identify possible threats by reviewing photos and video footage from surveillance equipment. This can speed up response times and aid in protecting military and civilian personnel.

The military has also employed computer vision for surveillance. Military personnel can keep an eye on important locations and gain intelligence by watching the video that is captured by cameras and drones. This can aid in spotting potential dangers and defending against assaults.

Computer vision has been utilized in the military for vehicle and pedestrian tracking in addition to target recognition and surveillance. Military personnel can trace the movements of people and vehicles by watching video footage, which helps to increase situational awareness and aid in navigation.

Overall, the military sector is more productive and effective because to the employment of computer vision.

V.DETECTING ANOMALIES

The ability to identify unusual patterns or deviations in data collected from a manufacturing system can be extremely useful for manufacturers. This is because identifying the cause of these anomalies can provide valuable insights about the system, such as the status of tools and machinery, and the quality of parts and products. By detecting and addressing these anomalies, manufacturers can prevent issues from escalating and ensure the smooth operation of their systems.

The increasing availability of low-cost sensors, wireless communication, and advanced computing, known collectively as the Industrial Internet of Things (IIoT), has made it possible for manufacturers to access large amounts of data from multiple sensors over time. This data, often in the form of multidimensional time series, can be used to more effectively monitor manufacturing systems. However, analyzing this data can also be more challenging, as the time series data may be long and high-dimensional, with each time step potentially containing acceleration and vibration signals from multiple axes and directions. Traditional anomaly detection methods may not be able to effectively analyze data of this length and complexity.

There are two main approaches for constructing reconstruction models for anomaly detection in time series data. The first approach involves using a portion of past data to predict future data. One such method involves using past time series data to predict multiple future time steps, and determining the probability of the predicted data being abnormal based on the distribution of the predictions. Another method involves using a predetermined length of past data to predict the signal of the next time step, and monitoring the reconstruction errors for anomalies. This approach has been applied to data from a jet engine to train a support vector machine model that can distinguish between anomalous and normal data.

The second approach involves using a series of neural networks as an encoder to summarize the time series data, and another series of neural networks as a decoder to reconstruct the data. This approach, known as the encoder-decoder structure, was first proposed for machine translation tasks and is generally considered to be more powerful than the first approach.

The proliferation of technologies such as smart sensors and the Internet of Things (IoT) has made it possible for manufacturers to gather a large amount of data from their systems, often in the form of a multidimensional time series with a high number of dimensions and a large number of time steps. This data can contain valuable information about failures and defects in the system. In this paper, we have proposed a new model for detecting anomalies in this type of time series data.

VI.SELF-DRIVING CARS

Self-driving cars, also referred to as autonomous or driverless cars, have been the focus of research and development by various organizations worldwide since the 1980s. Some examples of self-driving car research platforms from the past two decades include Navlab's mobile platform, the University of Pavia and University of Parma's car ARGO, and UBM's VaMoRs and VaMP vehicles.

The Defense Advanced Research Projects Agency (DARPA) held three competitions in the past decade to encourage the development of self-driving cars. The first competition, known as the DARPA Grand Challenge, took place in the Mojave Desert in 2004 and required self-driving cars to navigate a 142 mile course through desert trails within a 10 hour time limit. None of the cars were able to complete the course. The second DARPA Grand Challenge, held in 2005, required cars to navigate a 132 mile route through various terrains, including tunnels and sharp turns. There were 23 finalists in this competition, and four cars were able to complete the route within the allotted time. The Stanford University's car Stanley took first place, while the Carnegie Mellon University's cars Sandstorm and H1ghlander came in second and third place, respectively.

Self-driving cars (also known as autonomous or driverless cars) have been researched and developed by various organizations including universities, research centers, and car companies worldwide since the 1980s. The Society of Automotive Engineers (SAE) International has created a classification system to measure the level of autonomy of self-driving cars, ranging from level 0 (momentary human intervention is required) to level 5 (no human intervention is needed). This paper focuses on self-driving cars with an autonomy system at level 3 or higher as defined by the SAE. These cars typically have two main components: a perception system and a decision-making system. The perception system is typically made up of subsystems responsible for tasks such as localization, mapping, obstacle detection and tracking, and traffic signal recognition. The decision-making system is also divided into subsystems responsible for tasks such as route and path planning, behavior selection, motion planning, and obstacle avoidance and control.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. <https://emerj.com/ai-sector-overviews/military-applications-of-machine-vision-current-innovations/> (military)
2. <https://xane.ai/how-computer-vision-is-revolutionizing-combat-operations-in-the-armed-forces-globally/> (military)
3. <https://www.isahit.com/blog/10-computer-vision-applications-in-retail> (retail)
4. <https://viso.ai/applications/computer-vision-in-retail/> (retail)
5. <https://www.sciencedirect.com/science/article/abs/pii/S0278612520302211> (detecting anomalies)
6. <https://www.sciencedirect.com/science/article/abs/pii/S095741742030628X> (self-driving cars)
7. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**