Symbiosis

www.symbiosisonlinepublishing.com
ISSN Online: 2374-8141

Research Article | Nanoscience & Technology: Open Access | Open Access

Optimizing Computer Vision Algorithms: A MOORA (Multi-Objective Optimization by Ratio Analysis) Approach

Niranjan Mushke[1*], Bhanu Theja Musunuru[2] and Shanker Gangone[3]

[1, 2, 3]Incredible Software Solutions, Research and Development Division, Richardson, TX, 75080, USA

Received: October 16, 2019; Accepted: October 22, 2019; Published: December 26, 2019

***Corresponding author:** Niranjan Mushke, Incredible Software Solutions, Research and Development Division, Richardson, TX, 75080, USA,Email. niranjan.m369@gmail.com*

Abstract

Introduction: Computer Vision, a branch of artificial intelligence, empowers computers to interpret and understand visual data from the environment, much like human vision. This multidisciplinary field integrates principles from computer science, mathematics, and engineering to process and analyze images and videos. The main objective of computer vision is to automate visual tasks that humans perform, such as recognizing objects, segmenting images, detecting motion, and reconstructing scenes. Significant applications of computer vision include facial recognition, autonomous driving, medical image analysis, and industrial automation. In facial recognition, for example, algorithms can identify and verify people in images or videos. In autonomous vehicles, computer vision systems detect obstacles, road signs, and lane markings to ensure safe navigation. Computer vision employs a range of techniques from traditional image processing to advanced machine learning, especially deep learning. Convolutional Neural Networks (CNNs) have improved the precision and effectiveness of visual data interpretation, thereby revolutionizing the area. These networks detect patterns and features from extensive datasets, enabling more sophisticated and precise analysis. The rapid advancement in computer vision is fueled by increased computational power, the availability of large datasets, and algorithmic improvements. As technology progresses, computer vision is poised to become even more crucial across various industries, driving automation and unlocking new possibilities.

Research Significance: Computer vision research is crucial because it allows machines to comprehend and analyze visual information from their surroundings, leading to progress in multiple domains. It spurs innovation in autonomous vehicles, improving both safety and efficiency. In the healthcare sector, it enhances diagnostics and treatments through the analysis of medical images. Retail and manufacturing gain from automated inspection and inventory control. Moreover, computer vision strengthens security with sophisticated surveillance systems and supports augmented and virtual reality, enhancing user experiences. By empowering machines with the ability to "see," computer vision unlocks new opportunities for automation, accuracy, and intelligent decision-making across various industries.

Methodology: MOORA (Multi-Objective Optimization by Ratio Analysis) is a method used for evaluating alternatives based on multiple, often conflicting criteria. It helps in decision-making processes by allowing for the comparison of different performance ratios against reference alternatives. MOORA involves assigning weights to various criteria, measuring their significance, and sorting the alternatives accordingly. This technique aids in balancing multiple factors in complex decision-making scenarios. By systematically measuring trade-offs and considering multiple objectives simultaneously, MOORA provides a structured framework for achieving well-known and effective results.

Alternative: Convolutional Neural Networks (CNNs), Object Detection (YOLO), Image Segmentation (U-Net), Facial Recognition (DeepFace), Image Classification (ResNet), Optical Character Recognition (OCR), Video Analysis (OpenPose), 3D Reconstruction (Structure from Motion).

Evaluation Parameter: Accuracy (%), Processing Speed (FPS), Cost of Deployment ($), Energy Consumption (W), Data Requirements (GB), Maintenance Complexity.

Result: the result it is seen that Optical Character Recognition (OCR)is got the first rank whereas is the 3D Reconstruction (Structure from Motion) is having the lowest rank.

Keywords: Convolutional Neural Networks (CNNs), Object Detection (YOLO), Moora

Introduction

Advances in computer vision have transformed image data collection methods. Techniques like time lapse videos, automated counting via camera traps, contributions from citizen scientists, and data from aerial sensors have broadened our capabilities. Notably, automated detection algorithms are now widely used to identify large animals in images taken by high resolution commercial satellites and unmanned aerial vehicles. While commercial satellite imagery offers broad spatial coverage with detailed resolution, challenges such as atmospheric conditions, limited temporal coverage, and high costs persist. To detect animals within these images, researchers employ methods like pixel-based analysis, enhancing image contrast, and supervised classification using machine learning algorithms. Many applications focus on identifying groups of animals at breeding

 ***Corresponding author email:** niranjan.m369@gmail.com*

sites, capitalizing on the distinctive visual features and rich spatial context found in these environments. [1]The term "computer vision syndrome" (CVS) describes a group of symptoms brought on by extended computer use, impacting the eyes, muscles, and joints. As computers become increasingly integral to modern work environments, individuals may find themselves spending extended periods in front of screens, both in professional settings and elsewhere. Commonly reported symptoms among eye strain, exhaustion, burning or inflammation in the eyes, dryness, and discomfort are common among computer users, blurred vision at varying distances, headaches, as well as neck and shoulder discomfort. [2]Convolutional networks form the cornerstone of many computer vision solutions, with deep convolutional networks becoming increasingly important since their introduction in 2014. These networks consistently yield significant improvements across various metrics, despite the growing need for computational resources and larger datasets. They have demonstrated immediate quality enhancements, particularly when sufficient labeled data is available for training. Efficiently scaling networks while minimizing parameter counts and computational demands is crucial for applications such as mobile vision and big data visualizations. Techniques like Factored Curves and Aggressive Regularization are utilized to judiciously leverage additional computational resources in this pursuit. [3]Computer gaming has become an integral part of consumer electronics, captivating gamers with immersive experiences facilitated by a range of input devices like joysticks, buttons, trackballs, and motion-sensitive gloves. Increasingly, gamers seek interfaces enabling intuitive interaction through natural hand or body movements, driving the integration of computer vision technology. This advancement positions computer games at the forefront of embracing computer vision for enhanced user experiences. [4]Developing computer vision systems presents formidable challenges owing to the complexities of representing real-world scenarios and the extensive knowledge required. Machine learning (ML) presents notable advantages in tackling these obstacles by facilitating the creation of resilient and versatile vision systems across various settings. Nonetheless, leveraging ML for computer vision remains intricate and not entirely understood. Numerous unresolved issues persist, such as the challenge of transferring learned skills from one domain to another. Below are some of the persistent research challenges in this domain. [5]Computer vision has recently gained significant traction across diversescientific domains, such as industrial robotics and healthcare. Vision-based solutions are becoming more and more popular in the fields of architecture, engineering, construction, and facility management as a means of improving decision-making during the building phase. However, the chaotic and unpredictable nature of construction sites presents obstacles for effective monitoring. Substantial research has explored the potential of employing computer vision to assist in on-site management tasks. [6]Super-resolution (SR) reconstruction seeks to address the challenge of enhancing the clarity of images, especially in cases where the original images lack sufficient detail. This technique is utilized across various fields, with applications ranging from computer vision to other disciplines. One crucial aspect of solving the SR problem lies in assessing the transformation between images, which can either follow a straightforward parametric model or be tailored to specific scenes, necessitating point-by-point evaluation. Evaluating these transformations directly from images is essential and can be done both manually and automatically. [7]The increasing prevalence of urban surveillance cameras, especially in the field of computer vision. It explores advanced techniques in computer vision used for analyzing traffic videos, discussing current challenges and potential future research directions. These technologies show significant potential for Intelligent Transportation Systems (ITS), as decreasing hardware expenses make video analysis more accessible. ITS can leverage these systems for tasks like managing traffic congestion, enforcing traffic regulations, and tracking vehicles, which were traditionally done by humans. Common visual surveillance approaches such as background modeling and motion tracking are often used on highways to achieve effective vehicle detection and classification. [8]The rapid growth of deep learning algorithms has resulted in a significant revolution in computer vision. Two major elements have driven this transformation: the availability of large labeled picture datasets with millions of photos and considerable advancements in computer hardware that allow for faster computations. Prior to this paradigm change, models like LeCun's Neural Networks for Digit Recognition and Fukushima's Neocognitron centered computer vision research upon hierarchical representations. Although the machine learning and AI community held these early efforts in great regard, computer vision researchers in the 2000s tended to focus on other strategies. Scale-invariant features, bag-of-features, spatial pyramids, and similar techniques were the main methods used in these approaches. [9]My approach to computer vision is essentially the reverse of computer graphics. While computer graphics involves creating synthetic images by carefully modeling the world to simulate how images are formed, inverse computer graphics focuses on analyzing real images captured by cameras to build detailed geometric models. For example, a silhouette cone intersection technique was used to build the polyhedron in Figure 1 from several angles of a plastic horse on a turntable. This technique is essentially descriptive since it uses wide-angle stereo reconstruction. [10]In computer vision, images often contain outliers caused by noise, misalignment, or occlusion. Previous methods to improve PCA's robustness to these outliers are reviewed, and A novel approach to handle pixel-level outliers is presented, which makes use of an intra-sample outlier process. Robust M-estimation algorithm is introduced and the theory of Robust Principal Component Analysis (RPCA) is established for learning linear multivariate representations in high-dimensional data, such photographs. [11]The computer vision and robotics research community focuses on obtaining range data to support scene analysis, which enables the configurations are determined remotely (without human interaction)and spatial extents of three-dimensional object assemblages. This paper reviews various methods of generalized range finding, discussing their relevance and limitations within computer vision research. [12]Computer vision technology, although extensively utilized for inspection

across various industries, remains less common in aquaculture. The implementation of this technology in aquaculture faces substantial hurdles due to the challenging nature of the environment and the sensitivity of aquatic organisms. These organisms are prone to stress and move freely in conditions where lighting, visibility, and stability are often unpredictable. Furthermore, sensors need to operate effectively underwater or in moist environments. This review examines the current advancements and applications of computer vision technology in aquaculture, encompassing all production stages from hatcheries to harvest. [13]Deep learning's superior performance has led to its enormous rise in popularity recentlyin text, audio, and image processing tasks. This impressive capability has sparked interest in extending deep learning applications to more complex areas, such as3D handling of data. This article organizes the several methodologies that use deep learning for 3D data and looks at them. According to the reviewed literature, methods that use 2D projections of 3D data usually perform better than voxel-based (3D) deep learning models. But voxel-based models are able topotentially achieve superior results with the incorporation of additional layers and extensive data augmentation. [14] Regression analysis, a method utilized in computer vision to model relationships within noisy data, relies heavily on the least squares approach due to its widespread use and computational simplicity. While this method performs well when data errors adhere to a Gaussian distribution, its effectiveness declines in the presence of non-zero mean noise or outliersdata points that significantly stray from the expected pattern. Outliers can emerge from clutter, measurement inaccuracies, or sudden disturbances like impulse noise. In image analysis, for instance, boundary areas between different regions may generate outliers as samples from one region deviate when fitting models to neighboring areas. [15]

Material and Methods

Alternative: Convolutional Neural Networks (CNNs), Object Detection (YOLO), Image Segmentation (U-Net), Facial Recognition (DeepFace), Image Classification (ResNet), Optical Character Recognition (OCR), Video Analysis (OpenPose), 3D Reconstruction (Structure from Motion).

Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are specialized deep learning architectures designed primarily for processing structured grid data, such as images. A CNN consists of layers that automatically extract hierarchical features from images through convolutional layers, pooling layers, and fully connected layers. The convolutional layers apply filters to capture local patterns, while pooling layers reduce dimensionality, preserving essential features. This makes CNNs particularly effective for image classification, object detection, and more. They are widely used in applications like medical image analysis, autonomous vehicles, and facial recognition.

Object Detection (YOLO)

You Only Look Once (YOLO) is a state-of-the-art object detection algorithm that aims to identify and classify multiple objects within an image in real-time. Unlike traditional methods that apply a classifier to different regions of an image, YOLO treats object detection as a single regression problem, predicting bounding boxes and class probabilities directly from full images. This results in high-speed detection, making YOLO suitable for applications like surveillance, autonomous driving, and robotics, where real-time performance is critical.

Image Segmentation (U-Net)

U-Net is a convolutional network architecture primarily designed for biomedical image segmentation. It employs an encoder-decoder structure, where the encoder captures context through downsampling, and the decoder enables precise localization via upsampling. U-Net is particularly effective in tasks requiring pixel-wise classification, such as tumor detection in medical images or cell segmentation. Its architecture allows it to work well even with a limited number of training images, making it a go-to solution in medical image processing.

Facial Recognition (DeepFace)

DeepFace is a facial recognition system developed by Facebook that uses deep learning to analyze and recognize human faces. It employs a deep neural network to map facial features into a compact embedding space, allowing for efficient comparison between different faces. DeepFace achieves remarkable accuracy by leveraging a large dataset of labeled faces and utilizes techniques such as data augmentation to improve generalization. Applications include social media tagging, security systems, and access control.

Image Classification (ResNet)

Residual Networks (ResNet) revolutionized image classification by introducing skip connections, enabling the training of very deep networks without suffering from the vanishing gradient problem. By allowing gradients to flow through the network more effectively, ResNet can contain hundreds or even thousands of layers, achieving state-of-the-art performance on benchmarks like ImageNet. This architecture is widely used in various fields, including autonomous driving, medical diagnosis, and image retrieval systems.

Optical Character Recognition (OCR)

Optical Character Recognition (OCR) is a technology that converts different types of documents—such as scanned paper documents, PDFs, or images—into editable and searchable data. Modern OCR systems often utilize deep learning models to improve accuracy, leveraging CNNs and recurrent neural networks (RNNs) for character recognition. OCR applications span a wide range, including digitizing printed texts, automating data entry, and enabling accessibility features for visually impaired users.

Video Analysis (OpenPose)

OpenPose is a real-time multi-person keypoint detection library for human pose estimation. It identifies and tracks the positions of various body parts in videos, facilitating applications in sports analysis, animation, and augmented reality. By utilizing

a CNN-based architecture, OpenPose can process images to detect body, hand, face, and foot keypoints simultaneously, providing a comprehensive understanding of human activity in video feeds.

3D Reconstruction (Structure from Motion)

Structure from Motion (SfM) is a technique in computer vision for reconstructing 3D structures from 2D image sequences. It estimates the 3D geometry of a scene by analyzing how objects move between frames. SfM typically involves feature extraction, matching, and optimization to create a 3D point cloud representing the scene. This technology is crucial in applications such as virtual reality, drone mapping, and cultural heritage documentation.

Evaluation Parameter: Accuracy (%), Processing Speed (FPS), Cost of Deployment ($), Energy Consumption (W), Data Requirements (GB), Maintenance Complexity.

1. **Accuracy (%)**

Accuracy is a fundamental metric that measures how often a model makes correct predictions. In classification tasks, it is typically defined as the ratio of correctly predicted instances to the total instances. High accuracy is crucial, particularly in applications such as medical diagnosis or security systems, where false positives or negatives can have serious consequences. However, it's important to consider that accuracy alone can be misleading, especially in imbalanced datasets where one class significantly outnumbers others. Therefore, other metrics like precision, recall, and the F1-score should complement accuracy to provide a comprehensive evaluation.

2. **Processing Speed (FPS)**

Processing speed, often measured in frames per second (FPS), is critical for real-time applications, especially in scenarios like video analysis or autonomous driving. A system that can process 30 FPS may be suitable for real-time interaction, while 60 FPS is often required for smoother performance. Achieving high FPS can depend on various factors, including model complexity, hardware capabilities, and optimization techniques. For instance, lighter models or optimized frameworks can enhance processing speeds, making them suitable for edge computing applications where latency is a concern.

3. **Cost of Deployment ($)**

Cost of deployment encompasses all expenses related to implementing a machine learning model or computer vision system in a real-world environment. This includes hardware costs (servers, GPUs, sensors), software licenses, cloud service fees, and development costs (programming, testing). Cost considerations can significantly impact the choice of technology, especially for startups or small businesses. While high-performing models may require substantial upfront investment, understanding the total cost of ownership over time, including maintenance and operational costs, is essential for a comprehensive financial evaluation.

4. **Energy Consumption (W)**

Energy consumption is an increasingly important parameter, particularly in the context of sustainability and cost-efficiency. As machine learning models, especially deep learning models, can be computationally intensive, their energy requirements can be substantial. This is particularly relevant for edge devices and mobile applications where battery life is critical. Evaluating energy consumption involves measuring the watts used during inference or training. Reducing energy consumption can involve optimizing model architectures, using specialized hardware (like TPUs), or implementing energy-efficient algorithms that maintain performance while lowering power usage.

5. **Data Requirements (GB)**

Data requirements refer to the amount of data needed to effectively train and validate a model. Different models have varying data requirements based on their complexity and the tasks they are designed to perform. For instance, deep learning models typically require large datasets to generalize well, while traditional machine learning algorithms may perform adequately with smaller datasets. It's also essential to consider the quality of data; poorly labeled or noisy data can adversely affect model performance. Furthermore, data augmentation techniques can help mitigate data scarcity by artificially increasing the dataset size without the need for additional data collection.

6. **Maintenance Complexity**

Maintenance complexity addresses the ongoing efforts required to keep a system operational and effective. This includes regular updates, monitoring performance, retraining models with new data, and debugging. Systems that leverage machine learning often require more frequent maintenance compared to traditional software because they can drift over time due to changes in data patterns. The complexity can vary based on factors such as the model's architecture, the deployment environment (cloud vs. edge), and the need for human intervention. Simplifying maintenance may involve automating certain processes, using tools for continuous integration/continuous deployment (CI/CD), and ensuring documentation is thorough to facilitate troubleshooting.

MOORA (Multi-objective Optimization on the basis of Ratio Analysis)

The process of deriving ratios using this method is known as Multi-Objective Optimization on the Basis of Ratio Analysis (MOORA). This technique, which involves using dimensionless numbers, is a preliminary step towards optimization and can be applied to various districts in Lithuania. It effectively assesses differences in objectives among ten districts. While three districts are evaluated positively due to their prosperity, there is a marked contrast with less affluent districts. Additionally, labor migration across different districts plays a crucial role, potentially causing economic imbalances that may require corrective actions like automatic redistribution or measures to discourage migration. Alternatively, districts might consider strategies such as

commercialization and industrialization to foster development. [16]The MOORA (Multi-Objective Optimization by Ratio Analysis) approach is valuable in multi-objective optimization scenarios involving concurrent constraints or conflicting attributes. Its applications are broad, encompassing fields from manufacturing and process design to green choices and decision-making that involves trade-offs among competing objectives. MOORA's relevance is supported by Multiple Criteria Decision-Making (MCDM) techniques, serving three key roles. Firstly, it enhances traditional MCDM approaches by structuring factors in a sophisticated manner. Secondly, it addresses concerns about computation time often mentioned in MCDM literature. Lastly, MOORA is efficient, requiring minimal system resources in terms of processing time. Its practical use extends to educational settings, such as in the selection of university or college scholarships, where it helps prioritize students based on multiple attributes and can be applied in computerized assessments or other selection processes.[17]MOORA is a flexible and effective method for managing complex scenarios involving multiple attributes, standards, and conflicting goals. It systematically addresses multi-objective optimization challenges, often requiring trade-offs between competing criteria. This method is particularly useful for handling intricate and conflicting objectives, as it can adapt to various attributes with different levels of significance. Its clarity and versatility make it a practical choice for a range of situations. However, it's worth noting that MOORA may have limitations when confronted with certain disturbances.[18]The Multi-Objective Optimization on the Basis of Ratio Analysis (MOORA) is a robust method for tackling multi-objective optimization problems, accommodating various attributes and resolving conflicting factors. This technique proves valuable in scenarios with multiple standards or conflicting attributes, facilitating upgrades and enhancements. It excels in managing complexity arising from conflicting objectives and supply chain issues. MOORA is versatile, applicable in a range of areas such as assessments, provider selection, method design, and optimal solution determination. Furthermore, MOORA can be adapted to address identified failures by prioritizing them based on impact and severity. Through its extensions, MOORA integrates different analytical methods and credibility concepts to handle uncertainties associated with failures. The practical nature of MOORA ensures it provides realistic outcomes that support decision-makers. Comparisons with traditional methods underscore MOORA's effectiveness in identifying and resolving failures. [19]The MOORA method proves to be very effective when examined closely. For the researchers, it is clear that MOORA and the MOOSRA method are essential for selection processes, leveraging current data. From the earlier discussion, it's evident that both MOORA and MOOSRA meet all the criteria for decision-making challenges, making them highly reliable in production environments. When comparing them to the benefit-cost ratio, the rate's denominator charge indicates that these models provide a better overall financial performance.[20]Both the MOORA and MOOSRA methods are complex and distinctive, offering a thorough perspective on performance measurement techniques. They incorporate elements from the Rate Engine and Reference MOORA, blending attitudes with feature factors. We carried out extensive simulations for port planning, effectively setting goals, identifying types of substitutes, and assessing their importance. These methods are relevant to stakeholders, including national and local authorities and participating companies. In addressing production issues related to consumers, a clear focus on sovereignty is evident.[21]Officers, along with customers, frequently act as legal representatives, which can introduce some subjectivity and potential inaccuracies. To tackle problems related to the rating of CNC gadget devices, MOORA's decision-making framework is utilized. This framework incorporates complete information, known as a linguistic variable, which helps examiners manage ambiguous data. The article explores how multi-MOORA ranking is applied in different regions to provide a detailed overview, with the results of these rankings being summarized through evaluations. [22]

STEP 1: Design of decision matrix and weight matrix

For a MCDM problem consisting of m alternatives and n criteria, let $D = x_{ij}$ be a decision matrix, where $x_{ij} \in \mathbb{R}$

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}$$

The weight vector may be expressed as.

$$w_j = [w_1 \dots w_n], \text{ where } \sum_{j=1}^{n}(w_1 \dots w_n) = 1$$

$$n_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}$$

$$where\ i \in [1,m] and\ j \in [1,n]$$

STEP 3: Weighted normalized decision matrix

$$W_{nij} = w_j n_{ij}$$

STEP 4: Calculation of Performance value

The performance value of each alternative is calculated as

$$y_i = \sum_{j=1}^{g} N_{ij} - \sum_{j=g+1}^{n} N_{ij}$$

Where g is the number of benefit criteria and (n - g) is the cost criteria.

The alternatives are ranked from best to worst based on higher to lower y_i values.

Result and Discussion

Table 1.Computer Vision

	Accuracy (%)	Processing Speed (FPS)	Cost of Deployment ($)	Energy Consumption (W)	Data Requirements (GB)	Maintenance Complexity
Convolutional Neural Networks (CNNs)	92	30	2000	100	50	6
Object Detection (YOLO)	90	45	1500	120	40	5
Image Segmentation (U-Net)	95	25	1800	90	60	7
Facial Recognition (DeepFace)	98	35	2200	110	70	5
Image Classification (ResNet)	94	28	1700	95	50	6
Optical Character Recognition (OCR)	89	40	1200	80	30	4
Video Analysis (OpenPose)	91	30	1600	85	55	6
3D Reconstruction (Structure from Motion)	93	20	2500	150	100	7

Table 1 presents a comparison of various computer vision techniques across multiple performance metrics, including accuracy, processing speed, deployment cost, energy consumption, data requirements, and maintenance complexity.Convolutional Neural Networks (CNNs) achieve an accuracy of 92% with a processing speed of 30 FPS, costing $2000 to deploy, and consuming 100W of energy. Object Detection (YOLO) offers slightly lower accuracy (90%) but excels in speed (45 FPS) and lower deployment cost ($1500). In contrast, Image Segmentation (U-Net) provides high accuracy (95%) at a slower processing speed (25 FPS) and a deployment cost of $1800.Facial Recognition (DeepFace) boasts the highest accuracy (98%) but is the most expensive to deploy ($2200) and has moderate energy consumption. Image Classification (ResNet) balances performance with a respectable accuracy of 94% and a deployment cost of $1700. Optical Character Recognition (OCR), while less accurate (89%), remains cost-effective at $1200.Video Analysis (OpenPose) offers competitive performance metrics, and 3D Reconstruction (Structure from Motion) stands out for its high data requirements and deployment cost ($2500). Overall, the table highlights trade-offs in accuracy, speed, and resource needs among different computer vision approaches.

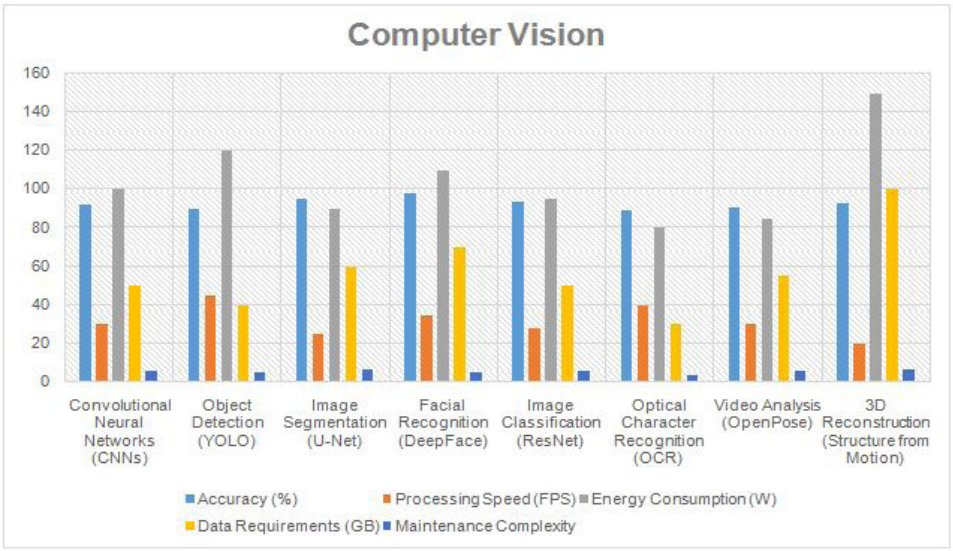

Figure 1: Computer Vision

Figure 1 presents a comprehensive comparison of various computer vision techniques across five key metrics: accuracy, processing speed, energy consumption, data requirements, and maintenance complexity.The graph showcases eight different computer vision approaches, including convolutional neural networks (CNNs), object detection, image segmentation, facial recognition, image classification, optical character recognition (OCR), video analysis, and 3D reconstruction.Accuracy (blue bars) is consistently high across all techniques, generally above 80%, indicating the reliability of these methods. Energy consumption (gray bars) varies significantly, with 3D reconstruction demanding the most power. Processing speeds (orange bars) are relatively lower compared to other metrics, suggesting these tasks are computationally intensive.Data requirements (yellow bars) fluctuate considerably between techniques, with 3D reconstruction and facial recognition needing the most data. Interestingly, maintenance complexity (light blue bars) is uniformly low across all methods, implying that once implemented, these systems are relatively easy to maintain.The graph effectively illustrates the trade-offs between performance and resource requirements for different computer vision applications, allowing for quick comparisons and informed decision-making when selecting an appropriate technique for specific use cases.

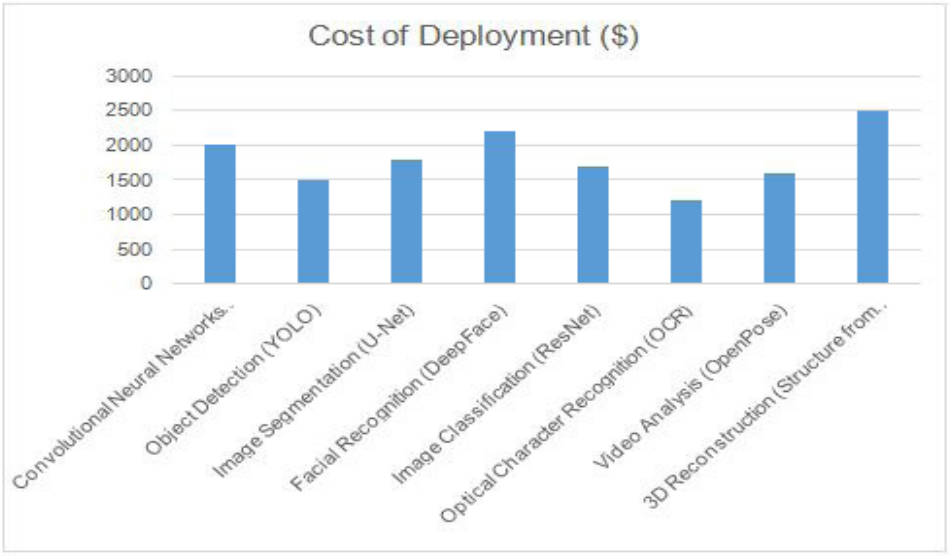

Figure 2: Cost of Deployment ($) (Computer Vision)

Figure 2 displays the deployment costs for various computer vision techniques. The graph compares eight different approaches, presenting their costs in dollars.3D Reconstruction stands out as the most expensive technique, with a deployment cost of about $2,500. This high cost likely reflects the complexity and resource-intensive nature of generating three-dimensional models from images. Facial Recognition (DeepFace) follows as the second most costly method, approaching $2,200. This could be due to the sophisticated algorithms and hardware required for accurate face detection and matching.Convolutional Neural Networks (CNNs) and Image Segmentation (U-Net) fall in the mid-range, both costing around $2,000 to deploy. These are fundamental techniques in computer vision with wide-ranging applications.At the lower end of the cost spectrum, we find Optical Character Recognition (OCR) as the least expensive option at approximately $1,200. This relative affordability might explain its widespread adoption in various industries.The graph provides valuable insights for decision-makers weighing the financial implications of implementing different computer vision technologies, helping to balance capability requirements against budget constraints.

Table 2. Divide & Sum

8464	900	4000000	10000	2500	36
8100	2025	2250000	14400	1600	25
9025	625	3240000	8100	3600	49
9604	1225	4840000	12100	4900	25
8836	784	2890000	9025	2500	36
7921	1600	1440000	6400	900	16
8281	900	2560000	7225	3025	36
8649	400	6250000	22500	10000	49
68880	8459	27470000	89750	29025	272

Table 2 outlines the performance metrics for various computer vision techniques in a "Divide & Sum" format, presenting their accuracy, processing speed, deployment costs, energy consumption, data requirements, and maintenance complexity in numerical values.Facial Recognition (DeepFace) stands out with the highest accuracy (9604) and significant processing speed (1225 FPS), indicating its effectiveness for real-time identification. However, it also incurs the highest deployment cost ($4,840,000) and energy consumption (12,100 W), highlighting the resource-intensive nature of this technology.Image Segmentation (U-Net) achieves a high accuracy of 9025 but has a lower processing speed (625 FPS) compared to others, reflecting a potential trade-off between precision and real-time performance. Object Detection (YOLO) offers a balanced approach with good accuracy (8100) and relatively high processing speed (2025 FPS) at a moderate cost.In contrast, Optical Character Recognition (OCR) is the most cost-effective ($1,440,000) and has the lowest maintenance complexity (16), making it suitable for budget-constrained applications.The totals at the bottom of the table indicate that these techniques collectively represent significant investments in deployment costs ($27,470,000) and energy consumption (89,750 W). Overall, this table highlights the varying trade-offs in performance, cost, and resource requirements across different computer vision technologies.

Table 3. Normalized Data

Normalized Data					
Accuracy (%)	Processing Speed (FPS)	Cost of Deployment ($)	Energy Consumption (W)	Data Requirements (GB)	Maintenance Complexity
0.3505	0.3262	0.3816	0.3338	0.2935	0.3638
0.3429	0.4893	0.2862	0.4006	0.2348	0.3032
0.3620	0.2718	0.3434	0.3004	0.3522	0.4244
0.3734	0.3805	0.4198	0.3672	0.4109	0.3032
0.3582	0.3044	0.3244	0.3171	0.2935	0.3638
0.3391	0.4349	0.2290	0.2670	0.1761	0.2425
0.3467	0.3262	0.3053	0.2837	0.3228	0.3638
0.3544	0.2175	0.4770	0.5007	0.5870	0.4244

Table 3 presents the normalized data for various computer vision techniques, allowing for a comparative analysis of their performance across key metrics. Each metric—accuracy, processing speed, deployment cost, energy consumption, data requirements, and maintenance complexity—has been scaled to facilitate easier comparison.Convolutional Neural Networks (CNNs) show moderate normalized scores across most metrics, with an accuracy of approximately 0.35 and a processing speed of 0.33. Object Detection (YOLO) excels in processing speed (0.49) while maintaining a competitive cost of deployment (0.29), indicating its efficiency in real-time applications.Image Segmentation (U-Net) achieves the highest normalized accuracy (0.36) but has a lower processing speed (0.27). Facial Recognition (DeepFace) presents a well-rounded profile with high accuracy (0.37) and acceptable processing speed (0.38), although it has the highest deployment cost (0.42).Optical Character Recognition (OCR) is notably cost-effective (0.23) and has the lowest energy consumption (0.27), making it suitable for low-resource environments. Video Analysis (OpenPose) and 3D Reconstruction (Structure from Motion) display varied strengths, with the latter having the highest data requirements and maintenance complexity. Overall, this normalized data highlights the trade-offs in efficiency and resource needs among different computer vision methods.

Table 4.Weight

Weight					
0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25
0.25	0.25	0.25	0.25	0.25	0.25

Table 4 presents a uniform weight distribution across multiple metrics in a computer vision context. Each metric, likely representing different performance aspects (e.g., accuracy, speed, cost), is assigned an equal weight of 0.25. This equal weighting indicates a balanced approach in evaluating the overall performance of various computer vision techniques.

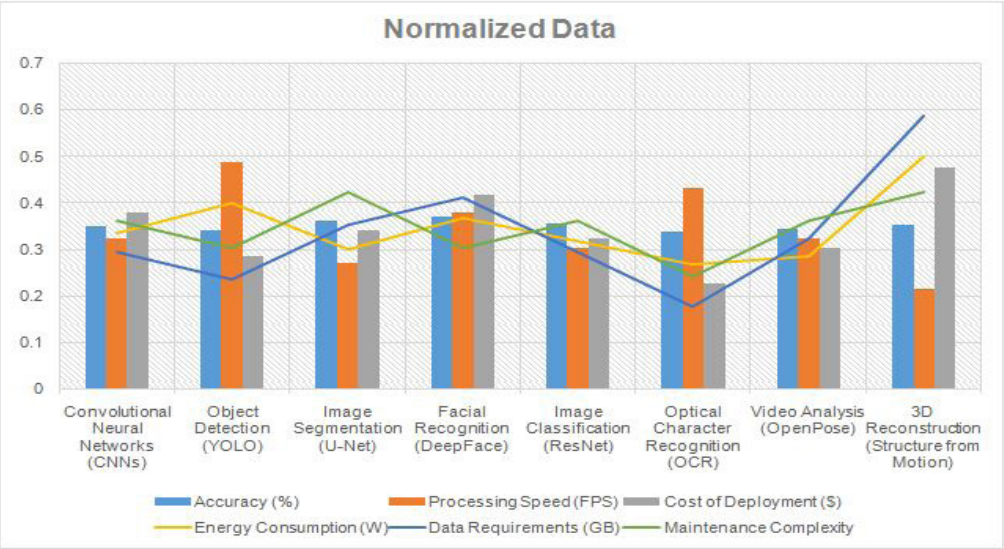

Figure 3: Normalized Data

Figure 3 presents a comparison of various machine learning models based on six key performance metrics: Accuracy, Processing Speed, Cost of Deployment, Energy Consumption, Data Requirements, and Maintenance Complexity. The data is normalized to a scale of 0 to 0.7, allowing for a direct comparison across different models.Convolutional models generally excel in accuracy but often require more data and have higher deployment costs.Object detection models offer a balance between accuracy and processing speed, making them suitable for real-time applications.Image and facial recognition models have moderate performance across most metrics, making them versatile choices for various tasks.Optical character recognition models are relatively efficient in terms of processing speed and energy consumption but may have limitations in accuracy for complex scenarios.Video analysis models typically require significant computational resources and data, but they can provide valuable insights for tasks like action recognition and anomaly detection.3D models are generally resource-intensive and may have limitations in terms of accuracy and speed, but they are essential for applications involving spatial data and object reconstruction.

Table 5. Weighted normalized decision matrix

Weighted normalized decision matrix					
0.087636	0.081546	0.095398	0.083449	0.073371	0.090951
0.085731	0.122319	0.071549	0.100139	0.058697	0.075792
0.090493	0.067955	0.085858	0.075104	0.088045	0.106109
0.093351	0.095137	0.104938	0.091794	0.102719	0.075792
0.089541	0.076109	0.081089	0.079277	0.073371	0.090951
0.084778	0.108728	0.057239	0.066759	0.044023	0.060634
0.086683	0.081546	0.076319	0.070932	0.080708	0.090951
0.088588	0.054364	0.119248	0.125174	0.146742	0.106109

Table5 presents a weighted normalized decision matrix for various computer vision techniques, where each metric's performance is adjusted based on its relative importance. This table facilitates a comparative analysis by assigning normalized values to key metrics: accuracy, processing speed, deployment cost, energy consumption, data requirements, and maintenance complexity.Facial Recognition (DeepFace) emerges as a strong candidate with the highest normalized values in accuracy (0.0934) and deployment cost (0.1049), indicating its superior performance in identification tasks despite its resource-intensive nature. Image Segmentation (U-Net) follows closely, with a commendable accuracy score (0.0905) but lower processing speed (0.0680), suggesting it excels in precision over real-time performance.Object Detection (YOLO) achieves a notable processing speed (0.1223), making it suitable for applications requiring

rapid analysis, albeit with a lower score in deployment cost (0.0715). Optical Character Recognition (OCR), while efficient in cost (0.0572), shows lower performance across other metrics, indicating its focus on specific use cases rather than overall versatility. Overall, this weighted matrix allows for a comprehensive evaluation, highlighting the strengths and weaknesses of each technique. It aids decision-makers in selecting the most suitable computer vision method based on prioritized criteria.

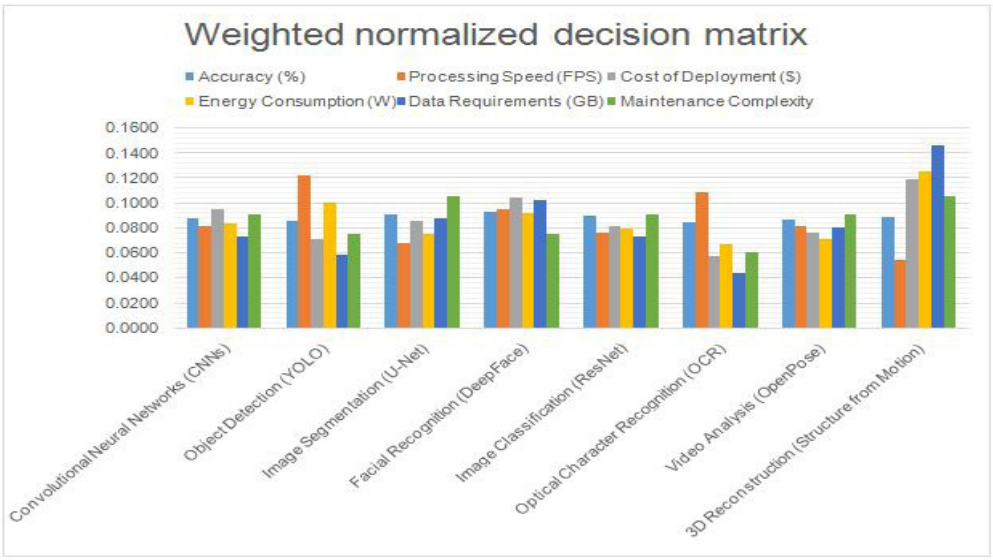

Figure 4: Weighted normalized decision matrix

Figure 4 presents a weighted normalized decision matrix for various computer vision techniques, comparing them across six key metrics: accuracy, processing speed, cost of deployment, energy consumption, data requirements, and maintenance complexity.The graph normalizes these metrics to a scale between 0 and 0.2, allowing for a fair comparison across different units of measurement. This normalization enables decision-makers to evaluate trade-offs between techniques more easily.Notable observations include:Object detection excels in processing speed but has higher energy consumption.3D reconstruction has the highest data requirements and deployment costs, but lower processing speed.Optical character recognition (OCR) shows a balance across metrics, with slightly higher processing speed.Facial recognition demonstrates high accuracy and data requirements, with moderate costs.Convolutional neural networks (CNNs) present a relatively balanced profile across all metrics.This visualization aids in multi-criteria decision-making, allowing stakeholders to prioritize specific attributes based on their project requirements. It highlights that no single technique dominates across all metrics, emphasizing the importance of considering use case-specific needs when selecting a computer vision approach.

Table 6. Assessment value & Rank

	Assessment value	Rank
Convolutional Neural Networks (CNNs)	0.016809	4
Object Detection (YOLO)	0.04497	2
Image Segmentation (U-Net)	-0.02495	7
Facial Recognition (DeepFace)	0.02312	3
Image Classification (ResNet)	0.00314	5
Optical Character Recognition (OCR)	0.079329	1
Video Analysis (OpenPose)	0.001957	6
3D Reconstruction (Structure from Motion)	-0.11583	8

Table 5 displays the assessment values and corresponding ranks for various computer vision techniques, providing a clear overview of their comparative performance. The assessment values are derived from a weighted analysis of multiple metrics, reflecting each technique's overall effectiveness.Optical Character Recognition (OCR) ranks highest with an assessment value of 0.079329, indicating its strong performance relative to other methods. This suggests OCR is particularly efficient and suitable for applications focused on

text recognition, likely due to its cost-effectiveness and lower resource demands.Following OCR, Object Detection (YOLO) secures the second rank with an assessment value of 0.04497, highlighting its quick processing capabilities and reasonable deployment costs, making it ideal for real-time applications. Facial Recognition (DeepFace) ranks third (0.02312), showcasing its robust accuracy but reflecting higher resource consumption.In contrast, 3D Reconstruction (Structure from Motion) has the lowest assessment value (-0.11583) and ranks eighth, indicating significant challenges in performance, likely due to high data requirements and deployment costs. Image Segmentation (U-Net) and Video Analysis (OpenPose) also perform poorly, with ranks of 7 and 6, respectively. Overall, this table aids in identifying the most effective techniques for specific applications, guiding decisions based on assessed values and ranks.

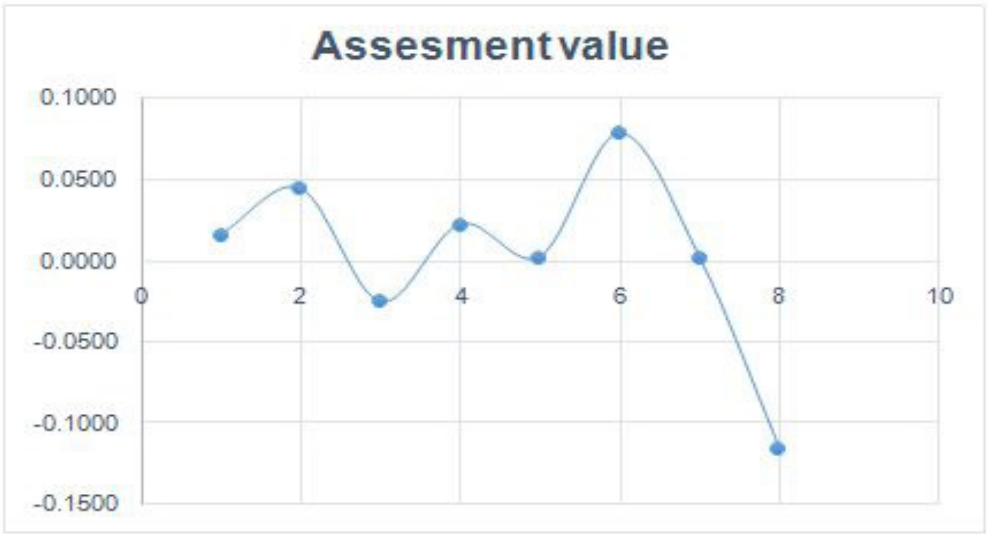

Figure 5: Assessment value

Figure 5 Assessment value Shows the Convolutional Neural Networks (CNNs)0.016809,Object Detection (YOLO)0.04497,Image Segmentation (U-Net)-0.02495,Facial Recognition (DeepFace)0.02312,Image Classification (ResNet)0.00314,Optical Character Recognition (OCR)0.079329,Video Analysis (OpenPose)0.001957,3D Reconstruction(Structure from Motion)-0.11583.

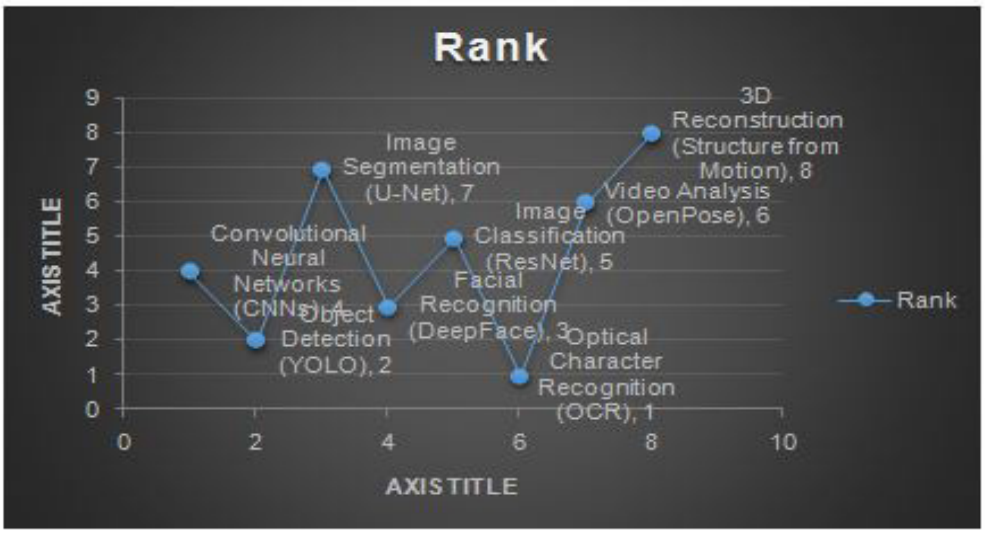

Figure 6: Assessment value

Figure 6 shows the Convolutional Neural Networks (CNNs) is in 4th Rank, Object Detection (YOLO) is in 2nd Rank, Image Segmentation (U-Net) is in 7th Rank, Facial Recognition (DeepFace) is in 3rd Rank, Image Classification (ResNet) is in 5th Rank, Optical Character Recognition (OCR) is in 1st Rank, Video Analysis (OpenPose) is in 6th Rank, 3D Reconstruction (Structure from Motion) is in 8th Rank

Conclusion

Computer vision, a branch of AI, aims to provide machines the same level of visual data understanding as humans. In order to obtain insightful information, digital photos must be collected, processed, and analyzed. The field of computer vision has made significant strides in recent years, primarily due to developments in deep learning and machine learning, as well as the exponential rise in processing power and data availability. At its essence, computer vision aims to automate tasks traditionally performed by the human visual system, such as creation, segmentation, object detection, and image recognition. In this field, the introduction of convolutional neural networks (CNNs), a kind of deep learning model, has proved essential, as they can learn hierarchical features from raw pixel data, significantly enhancing accuracy and efficiency in complex image-related tasks. The applications of computer vision span across diverse fields. In medicine, it aids in disease diagnosis through precise analysis of medical imaging like X-rays and MRI scans. In autonomous vehicles, It is essential for activities like object identification, lane detection, and driver monitoring since it the development of safer self-driving technology. Similarly, in retail, computer vision facilitates inventory management, automated checkout, and personalized shopping experiences. Additionally, it enhances surveillance and security measures through advanced facial recognition and behavior analysis systems. Despite its advancements, computer vision faces challenges, notably the requirement for extensive labeled data for model training, which can be resource-intensive. Furthermore, deep learning models are often perceived as black boxes, lacking interpretability, and may struggle with robustness and generalization across diverse environments. Looking forward, the future of computer vision appears promising. Ongoing research aims to overcome existing limitations and explore new avenues. Techniques like transfer learning and unsupervised learning hold potential to reduce reliance on labeled data. Moreover, efforts towards explainable AI seek to enhance transparency and understanding in computer vision models. Integration with complementary AI technologies like natural language processing and robotics is expected to further broaden its applications. In summary, computer vision serves as a fundamental aspect of modern AI, revolutionizing how machines perceive and interact with visual information. Its impact spans various sectors, and continual advancements promise to push boundaries further. Addressing challenges such as data dependency and interpretability will lead to more reliable and versatile computer vision technologies.

References

1. Weinstein, Ben G. "A computer vision for animal ecology." Journal of Animal Ecology 87, no. 3 (2018): 533-545.
2. Blehm, Clayton, Seema Vishnu, Ashbala Khattak, Shrabanee Mitra, and Richard W. Yee. "Computer vision syndrome: a review." Survey of ophthalmology 50, no. 3 (2005): 253-262.
3. Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. "Rethinking the inception architecture for computer vision." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826. 2016.
4. Freeman, William T., Ken-ichi Tanaka, Jun Ohta, and Kazuo Kyuma. "Computer vision for computer games." In Proceedings of the second international conference on automatic face and gesture recognition, pp. 100-105. IEEE, 1996.
5. Sebe, Nicu. Machine learning in computer vision. Vol. 29. Springer Science & Business Media, 2005.
6. Xu, Shuyuan, Jun Wang, Wenchi Shou, Tuan Ngo, Abdul-Manan Sadick, and Xiangyu Wang. "Computer vision techniques in construction: a critical review." Archives of Computational Methods in Engineering 28 (2021): 3383-3397.
7. Capel, David, and Andrew Zisserman. "Computer vision applied to super resolution." IEEE Signal Processing Magazine 20, no. 3 (2003): 75-86.
8. Buch, Norbert, Sergio A. Velastin, and James Orwell. "A review of computer vision techniques for the analysis of urban traffic." IEEE Transactions on intelligent transportation systems 12, no. 3 (2011): 920-939.
9. Ponti, Moacir Antonelli, Leonardo Sampaio Ferraz Ribeiro, Tiago Santana Nazare, Tu Bui, and John Collomosse. "Everything you wanted to know about deep learning for computer vision but were afraid to ask." In 2017 30th SIBGRAPI conference on graphics, patterns and images tutorials (SIBGRAPI-T), pp. 17-41. IEEE, 2017.
10. Baumgart, Bruce G. "A polyhedron representation for computer vision." In Proceedings of the May 19-22, 1975, national computer conference and exposition, pp. 589-596. 1975.
11. De la Torre, Fernando, and Michael J. Black. "Robust principal component analysis for computer vision." In Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, vol. 1, pp. 362-369. IEEE, 2001.
12. Jarvis, Ray A. "A perspective on range finding techniques for computer vision." IEEE Transactions on Pattern Analysis and Machine Intelligence 2 (1983): 122-139.
13. Zion, Boaz. "The use of computer vision technologies in aquaculture–a review." Computers and electronics in agriculture 88 (2012): 125-132.
14. Ioannidou, Anastasia, Elisavet Chatzilari, Spiros Nikolopoulos, and Ioannis Kompatsiaris. "Deep learning advances in computer vision with 3d data: A survey." ACM computing surveys (CSUR) 50, no. 2 (2017): 1-38.
15. Meer, Peter, Doron Mintz, Azriel Rosenfeld, and Dong Yoon Kim. "Robust regression methods for computer vision: A review." International journal of computer vision 6 (1991): 59-70.
16. Jahan, Ali, Faizal Mustapha, Md Yusof Ismail, S. M. Sapuan, and Marjan Bahraminasab. "A comprehensive VIKOR method for material selection." Materials & Design 32, no. 3 (2011): 1215-1221.
17. Sayadi, Mohammad Kazem, Majeed Heydari, and Kamran Shahanaghi. "Extension of VIKOR method for decision making problem with interval numbers." Applied Mathematical Modelling 33, no. 5 (2009): 2257-2262.
18. Chatterjee, Prasenjit, and Shankar Chakraborty. "A comparative analysis of VIKOR method and its variants." Decision Science Letters 5, no. 4 (2016): 469-486.
19. Liu, Hu-Chen, Jian-Xin You, Xiao-Yue You, and Meng-Meng Shan. "A novel approach for failure mode and effects analysis using

combination weighting and fuzzy VIKOR method." Applied soft computing 28 (2015): 579-588.

20. Liao, Huchang, Zeshui Xu, and Xiao-Jun Zeng. "Hesitant fuzzy linguistic VIKOR method and its application in qualitative multiple criteria decision making." IEEE Transactions on Fuzzy Systems 23, no. 5 (2014): 1343-1355.
21. Zhang, Nian, and Guiwu Wei. "Extension of VIKOR method for decision making problem based on hesitant fuzzy set." Applied Mathematical Modelling 37, no. 7 (2013): 4938-4947.
22. Chang, Chia-Ling. "A modified VIKOR method for multiple criteria analysis." Environmental monitoring and assessment 168, no. 1 (2010): 339-344.
23. Kim, Jong Hyen, and Byeong Seok Ahn. "Extended VIKOR method using incomplete criteria weights." Expert Systems with Applications 126 (2019): 124-132.