Computer Vision and Image Processing the Challenges and Opportunities for new technologies approach: A paper review

Reem M. Hussien 1*, Karrar Q. Al-Jubouri 2, Mohaimen Al gburi 3, Al Gburi Hussein Qahtan 4, Al Hamami Duaa Jaafar 4

1 Department of Civil Engineering, Al-Farabi University College, Baghdad, Iraq.
2 Department of Production Engineering and Metallurgy, University of Technology, Iraq.
3 Dep of Electrical and Computer Engineering, Altınbaş University, Istanbul, Turkey.
4 Administrative Affairs, ID dep, Ministry of Education, Baghdad, Iraq.

karrar.q.khalaf@uotechnology.edu.iq, mohaimen.algburi@ogr.altinbas.edu.tr *, hussein.kahtan@gmail.com

Abstract. Digital image processing has numerous applications in many sectors of the world. It expands from initial information registration into methods and thoughts combining pattern recognition, computer vision, and machine learning. The wide utilization has pulled in many researchers to integrate with a range of related specializations. This work gives a study of the latest development and theoretical ideas clarifying an improvement of computer vision particularly with pattern recognition and image processing, utilizing various regions of their field implementation. Digital image processing causes researchers to analyze images to get important data and comprehend information. It utilized a technique of multi-range implementation and huge information analysis. This work aims to focus on the latest studies related to image processing, pattern recognition, and computer vision. In this paper, Computer vision standard has been categorized into groups. For example, pattern recognition, image processing, and AI. Additionally, we give a short clarification of the forward-thinking data about the methods and their realization. This survey is restricted to computer vision, and more research can include detecting the conduct and properties of the object including people actions.

1. Introduction
Digital image processing has been extended for a numerous field going from initial information registration into the techniques of the computer graphics, object detection, and data interpretation [1]-[3]. It has a blend of ideas, methods, and thoughts from computer vision, object recognition, image processing, and AI. Almost the majority of the assignments in computer vision are regarding the way to get data on events or adjectives, from object detection and focus to get features [4]. The techniques utilized to solve issues in image processing depended on the application field and the type of information which has analysed. First of all, Image processing is divided into computer vision [6]. The bottom line of the vision system process is image recognition. Advancement of this area via achieved adjust the capability of human vision in getting data [7]. Computer vision is an order to extract data from the image, instead of image processing. Develop of the computer vision relies upon the image processing system, regardless of whether about pixel enhancement or item detection [8]. There is interference with pattern recognition on essential methods, and a few authors utilize the two terms
interchangeably. The primary motivation behind developing Computer vision is to make paradigms to get information and data from images, on the other hand, Image Processing is an achievement in actualizing computational changes for images, for example, contrast, sharpening, [9] etc. It additionally has a comparable importance and once in a while interferes within Human and Computer Interaction. HCI is then evolved as a different control (It is considered a multidisciplinary field) which talks about the mutual relations between computer-human interceded through technology improvement including the human side, the point of deciphering spatial information for example information ordered by more than one measurement [10]. In any case, vision systems cannot be relied upon to imitate quite the human eye. It is because of a computer vision framework has finite realization and task, compared to the natural eye [5]. Despite the fact that numerous researchers have suggested a wide zone of computer vision manner to similar the natural eye, but, much of the time, there are some challenging functions faced computer vision systems [11]. The core difficulties in their performance is the quality of an algorithm, the affectability of the parameters, and the precision of the outcomes. For the most part, the performance assessment includes calculating a several of the essential conducts of an algorithm to accomplish exactness, quality, or extendibility to lead and observe system. Because the efficiency of vision systems relies upon the use system structure, there are efforts suggested by numerous researchers to extend and arranged computer vision into numerous fields and particular applications, for example, robotics, remote sensing, automation on the mechanical production system [12]-[13], and others.

2. Literature Review

Computer vision depends on optical sensors and algorithms to animate human visualization so as to extract important data from the items [13]. Contrasted with traditional ways that take quite a while and require complicated research centers, Computer vision has developed into a part of image processing and artificial intelligence. (D.E. Goldberg) in 1989, issued a monograph (Gonzalez et al., 2016). The book sums up the core outcomes of genetic algorithm study. The applications of Genetic algorithms and usage have been exhaustively and efficiently discussed. Around the same time, the researcher Koza from Stanford University in USA inventively suggested a genetic programming technique that utilizes various computer programs to manifest issues dependent on the standard of normal selection (Wang et al., 2015). It effectively solved numerous issues. In 1995, Koza issued research, which developed the exploration of genetic programming and made the robotics show another state (Wang et al., 2017). The in-progress advancement of image processing innovation have given the probability to make a system to identify pixels. The advancement is summed up in table 1.

| Table 1. The role of researchers in the evolution of image processing |
|---------------------------------------------------------------|
| Researchers          | interest                                      | Definition                                                                 |
|----------------------|-----------------------------------------------|---------------------------------------------------------------------------|
| Savioja et al., [14] | image processing in an audio system          | Audio data saved in the media, for example, phonograph recording can be recreated, with no least contact, by computing the shape of the section utilizing digital image processing. |
| Xiaogang Wang et al., [15] | object detection via (CNN) | Find a solution to border detection issues in image processing regions by ANN in in-depth learning. |
| Liu et al., [16]     | computer vision                               | Examination of pixel quality with various technique rapidly and dependably. Invert arrange (top-down). |

Pattern recognition works to identify objects through algorithms applied to the image to obtain an improved image and understanding of the image [17]. Studies and technologies related to this field
aim to improve and develop algorithms for collecting and processing information then extracting features in the digital image, musical phrase, or syllable representing a word or even computer text \[10\], shown in fig. (1, 2).

![Figure 1. Pattern Recognition Processing](image1)

![Figure 2. Pattern Recognition System](image2)

The computer vision systems have comprised two fundamental gatherings, for example, a three-dimensional examination and pixel enhanced. A 3D morphological examination is a basic hypothesis for pattern recognition and computer graphics processing, while pixel enhance is connected to the description of morphological operation, fig 3.

![Figure 3. Morphological processing in image](image3)
Also, the process must take place on huge data group, coverage several layers of engineering synthesis. Thus, computing algorithms precise and effectual to extricate the significant quantitative data are critical to comprehend the compound color clusters as all [17]. The collection between some AI techniques with morphological analysis can outcome the best achieve via computing algorithms. The algorithms are genetic algorithms, artificial neural networks, and fuzzy logic [18]-[20].

Image Segmentation is a process of dividing the image into parts or a group of pixels through specific algorithms [21]. The aims to change its representation and make it simpler and easier for analytical operations. If humans can naturally separate objects in an image, it is thanks to high-level knowledge. Developing high-level segmentation algorithms (each region is a semantic object) is still one of the majority popular research topics in image processing [22].

Image Segmentation is widely used, the most famous being in a self-driving cars, as shown in Fig. 4, people were divided into red, cars in blue, the road in purple, and others. These sections facilitate the processing and analysis of objects in the image or video.

![Image Segmentation](image.jpg)

**Figure 4. Image segmentation**

In many applications of computer vision, boundary/edge identification and image division is significant in the item recognition and realization.

Object Detection is the procedure of recognizing cases of physical objects of a specific kind [23], (for example, people, planes, or animals) in video or images (Fig 5). A popular method for object discovery system involves creating a wide variety of filter windows in a sequence categorized utilizing CNN features. For instance, the technique portrayed in utilizes a choosy search to obtain items suggestion, educe CNN features for every proposition, and afterward feeds features to SVM to determine if a window contains the item or not. Countless works depend on the idea of CNN features with Regions suggestion [24].

There are countless manner of attempting to increase optimize Regions performance using CNN and path-planning approaches, [25]-[27] some of them succeed in discovering the location of the object but it is not possible to regularly define the accurate location of the item. To achieve this, such techniques frequently follow a joint item recognition—semantic division approach, ordinarily achieving great outcomes.

However, there is a few number of attempts to discover the object utilizing deep models. For instance, [28] suggest a coarse item determination technique dependent on a saliency component related to a DBN for object identification in remote detecting images, while takes advantage of the stacked auto encoders to discover multiple organs in medical images, and it takes advantage of the stacked automatic encoders directed to detect outstanding video-based objects [29]-[31].
Pixel improvement is helpful for object identification, division, and recognition. In limit edge-based methods, a border identifier is utilized to locate the border of an item. This technique depends on the sharpness of pixels that will change quickly on the edge of two areas. For shading division, border locate is done to each RGB colour channel [32]. This creates borders that can be integrated to create the definitive edge image. Pixels are assembled based on uniformity standards. For example, growth and separation techniques where pixels are assembled into bigger regions dependent on predefined standards. This procedure began from a collection of fundamental points and expansion into large regions by including each close pixel from the above beginning stages that have comparable features, for example, gray values or gray color. The procedure of combining and separation pixels start by segmenting the image into various areas and afterward merging the regions to achieve predetermined standards. The technique based on Region has two significant shortcomings. Firstly, the mechanism of separation and combination of area evolution based on global criteria models. Secondly, the procedure of regional growth relies upon the initial pixel and segments which affects the performance of the detection of the object [33]-[35], Figure 6.
3. Conclusion

Computer Vision and Pattern Recognition have been related to machine learning and Image processing. Computer vision applications are a wide area, interfere with many disciplines which have closely related to the digital image processing system in image analysis and information extraction to perform tasks accurately, there are several areas that will be developed with the help of the vision systems, like robots, self-driving vehicles, and items detection, etc. Researchers in the field of object detection and pattern recognition that are a branch of computer vision can develop this field by improving algorithms to interpret the image and obtain features to predict properties and object behaviour including natural events and human.

4. Acknowledgments

The author, Al Gburi Hussein Qahtan, would like to express special thanks to Mr. Ali Saud Abdul Majeed and Mr. Imran Musa Rada in the Ministry of Education– Iraq for their support for this research.

5. References

[1] Kulkarni, A. S., & Shinde, S. B. (2017, April). A review paper on monitoring driver distraction in real time using computer vision system. In Electrical, Instrumentation and Communication Engineering (ICEICE), 2017 IEEE International Conf. on (pp. 1-4). IEEE.

[2] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.

[3] C. Wang, Z. Zhang, and Z. Xi, “A human body based on sift-neural network algorithm attitude recognition method,” Journal of Medical Imaging and Health Informatics, vol. 10, no. 1, pp. 129–133, 2020.

[4] J. Yim, D. Joo, J. Bae, and J. Kim, “A gift from knowledge distillation: fast optimization, network minimization and transfer learning,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4133–4141, Honolulu, HI, USA, July 2017.

[5] G. Xie, J. Wang, T. Zhang, J. Lai, R. Hong, and G.-J. Qi, “Interleaved structured sparse convolutional neural networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8847–8856, Salt Lake City, UT, USA, June 2018.

[6] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. “Fully convolutional networks for semantic segmentation.” In Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition. pp. 3431-3440. 2015.

[7] X. Liu, K. N. Khan, Q. Farooq, Y. Hao, and M. S. Arshad, “Obstacle avoidance through gesture recognition: business advancement potential in robot navigation socio-technology,” Robotica, vol. 37, no. 10, pp. 1663–1676, 2019.

[8] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, “ChestX-Ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases,” in Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3462–3471, Honolulu, HI, May 2017.

[9] G. Xie, J. Wang, T. Zhang, J. Lai, R. Hong, and G.-J. Qi, “Interleaved structured sparse convolutional neural networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8847–8856, Salt Lake City, UT, USA, June 2018.

[10] Y. Alshawabkeh, “Linear feature extraction from point cloud using color information,” Heritage Science, vol. 8, no. 1, pp. 1–13, 2020.

[11] J. Li, C. Xia, and X. Chen, “A benchmark dataset and saliency-guided stacked autoencoders for video-based salient object detection,” IEEE Transactions on Image Processing, 2017.

[12] W. Ouyang, X. Zeng, X. Wang et al., “DeepID-Net: Object Detection with Deformable Part Based Convolutional Neural Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 7, pp. 1320–1334, 2017.
[13] Z. Qin, Y. Zhang, S. Meng, Z. Qin, and K.-K. R. Choo, “Imaging and fusing time series for wearable sensor-based human activity recognition,” Information Fusion, vol. 53, pp. 80–87, 2020.

[14] Savioja, Lauri, Akio Ando, Ramani Duraiswami, Emanuel AP Habets, and Sascha Spors. "Introduction to the issue on spatial audio." IEEE Journal of Selected Topics in Signal Processing 9, no. 5 (2015): 767-769.

[15] Yang, Wei, Wanli Ouyang, Hongsheng Li, and Xiaogang Wang. "End-to-end learning of deformable mixture of parts and deep convolutional neural networks for human pose estimation." In Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition, pp. 3073-3082. 2016.

[16] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In Proceedings of the IEEE conf. on computer vision and pattern recognition, pp. 1-9. 2015.

[17] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In Proceedings of the IEEE conf. on computer vision and pattern recognition, pp. 1-9. 2015.

[18] S. Ghosh, S. K. K. Srinivasa, P. Amon, A. Hutter, and A. Kaup, “Deep network pruning for object detection,” in Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), pp. 3915–3919, IEEE, Taipei, Taiwan, September 2019.

[19] W. Ouyang, X.Wang, X. Zeng et al., “DeepID-Net: Deformable deep convolutional neural networks for object detection,” in Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition, CVPR 2015, pp. 2403-2412, USA, June 2015.

[20] D. Jiang, G. Li, Y. Sun, J. Kong, and B. Tao, “Gesture recognition based on skeletonization algorithm and CNN with ASL database,” Multimedia Tools and Applications, vol. 78, no. 21, pp. 29953–29970, 2019.

[21] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: inverted residuals and linear bottlenecks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4510–4520, Salt Lake City, UT, USA, June 2018.

[22] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going deeper with convolutions." In Proceedings of the IEEE conf. on computer vision and pattern recognition, pp. 1-9. 2015.

[23] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.

[24] B. Heo, M. Lee, S. Yun, and Y. C. Jin, “Improving knowledge distillation with supporting adversarial samples,” in Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, February 2019.

[25] W. Diao, X. Sun, X. Zheng, F. Dou, H. Wang, and K. Fu, “Efficient Saliency-Based Object Detection in Remote Sensing Images Using Deep Belief Networks,” IEEE Geoscience and Remote Sensing Letters, vol. 13, no. 2, pp. 137–141, 2016.

[26] H. Al Gburi, and A. Emad, “Path-Planning Dynamic 3D Space Using Modified A* Algorithm,” IOP Conference Series: Materials Science and Engineering, 2020, 928(3), 032016, DOI: 10.1088/1757-899X/928/3/032016.

[27] A. Emad, H. Al Gburi, and D. Jaafar, “Securing software defined network transactions using visual cryptography in steganography,” Periodicals of Engineering and Natural Sciences, 2020, 8(4), pp. 2405–2416. DOI: 10.21533/pen.v8i4.1737.

[28] J. Frankle and M. Carbin, “The lottery ticket hypothesis: finding sparse, trainable neural networks,” in Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, May 2019.
[29] H.-C. Shin, M. R. Orton, D. J. Collins, S. J. Doran, and M. O. Leach, “Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1930–1943, 2013.

[30] J. Li, C. Xia, and X. Chen, “A benchmark dataset and saliency-guided stacked autoencoders for video-based salient object detection,” IEEE Transactions on Image Processing, 2017.

[31] P. Molchanov, A. Mallya, S. Tyree, I. Frosio, and J. Kautz, “Importance estimation for neural network pruning,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 11264–11272, Long Beach, CA, USA, June 2019.

[32] J. Cui, C. Min, and D. Feng, “Research on pose estimation for stereo vision measurement system by an improved method: uncertainty weighted stereopsis pose solution method based on projection vector,” Optics Express, vol. 28, no. 4, pp. 5470–5491, 2020.

[33] K. Ma, Q. Wu, Z. Wang, Z. Duanmu, H. Yong, H. Li, and L. Zhang. Group mad competition - a new methodology to compare objective image quality models. In The IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), June 2016. 7.

[34] D. Jaafar, H. Al Gburi, and A. Emad, “Path-planning in 3D space using butterfly optimization algorithm,” Periodicals of Engineering and Natural Sciences, 2021, 9(2), pp. 365–374.

[35] M. K. Quweider, J. D. Scargle, and B. Jackson, “Grey level reduction for segmentation, thresholding and binarisation of images based on optimal partitioning on an interval,” IET Image Processing, vol. 1, no. 2, pp. 103–111, 2007.