IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A Survey Paper On Object Detection And Localization Methods InImage Processing

[1] Katroth Balakrishna Maruthiram, [2] Dr.G.Venkata Rami Reddy,

[1]Assistant Professor of CSE, [2]Professor of IT,
[1]Department of IT, JNTUCEST, JNTUH
[2]Department of IT, JNTUCEST, JNTUH
[1]Jawaharlal Nehru Technological University Hyderabad, India
[2]Jawaharlal Nehru Technological University Hyderabad, India

***Abstract*—** over the past few decades, advances in information technology have significantly transformed how we manage data and information, particularly in terms of data acquisition, processing, and predictions. This interdisciplinary effort has involved the use of image processing techniques and the implementation of AI-based models. The latest technological innovations have made it possible for researchers to conduct computational experiments that would have been impractical using conventional methods. This survey paper examines various methodologies for object detection, providing a systematic analysis of existing techniques with a detailed presentation. The study also includes an evaluation of the strengths and weaknesses of current methods and outlines future research opportunities in this field.

***Keywords*—** Object Recognition, Object Detection, Image Processing, AI-Artificial Intelligence.

I. INTRODUCTION

The acquisition, storage, and processing of digital images have become increasingly important in our daily lives. The advancements in digital image-processing technologies have completely revolutionized how we digitize and share images across various web and mobile platforms. Now, we can instantly share photos online, send and receive multimedia, exchange audio and video messages, and even stream live videos worldwide. These innovations have been made possible by advancements in both software and hardware supporting the digital imaging field.

Analyzing large volumes of data is essential in today's world, and advancements in Big Data and Cloud Computing technologies have significantly improved data storage and analysis. Platforms such as AWS, Azure, and Google Cloud enable seamless transfer of data from on-premise servers to the cloud, making it accessible to authorized users. Cloud-stored data is secure due to the maintenance of multiple backup copies by the cloud vendors.

The term Object recognition is generally used to refer to computer vision tasks that involve identifying objects in digital images, while image classification involves predicting the class of an object, and object localization involves identifying the location of an object in an image and drawing a bounding box around it. Object detection, the main focus of this survey paper, often combines both classification and localization tasks in an image. Examples of object detection include face detection and pedestrian detection. Object detection has become a popular research area for scholars, aiming to replicate human intelligence in quickly identifying objects within images or videos. These methods primarily rely on machine learning algorithms to locate objects, draw bounding boxes around them, and determine their class. Object detection has various applications including face detection, OCR recognition, autonomous vehicles, video surveillance, pedestrian detection, and medical imaging.

II. REVIEW OF LITERATURE

In order to give a comprehensive technical study of the object identification techniques, nearly 100 publications were downloaded as part of the survey review from a variety of digital libraries, including IEEE Xplore, Science Direct, Google Scholar, ACM, and many more. The 24 best papers for the review have been chosen and are organized in this portion of the article in a methodical manner after careful examination of the paper's title, abstract, introduction, experiment, and future scope.

The literature study examines prior efforts in the domain of object identification research and development, utilizing several machine learning and deep neural network techniques. Several academics in this field have made important contributions to the literature survey discussion.

The authors of this study have attempted to tackle the issue of cutting down on the time and effort required to physically inspect the fruits' quality [1].

The authors have designed and developed a model for fruit assessment using a variety of library functions and model elements. The color, size, texture, form, and weight of an ideal fruit are first captured by the system. After the photos are taken, comparisons are done, and the results are displayed as output. When the experiment's findings were compared to earlier research, it was discovered that there had been an improvement in accuracy and a decrease in effort.

The authors of this study have presented a novel method for object detection through visual cues. The effectiveness of these strategies in distinguishing and detecting the right object from a plethora of other objects was demonstrated through experiments. This involved determining the precise characteristics and accurate properties of the geometric shapes. When there are numerous other things in the background, the strategies are quite helpful in distinguishing the correct objects [2].

The authors of this study have suggested a method to deal With the difficulties caused by visual agnosia among patients. People with this kind of disease are unable to distinguish between two-dimensional (2-D) visuals. The image sizes were compared to real-life photos in order to determine whether the authors' methodology was congruent or not [3].

The authors of this work have found that the dorsal visual cortex primarily mediates object size. The authors combined several photos using a Matlab tool, which logically led to the production of a high-resolution output. To overlap input photographs that were taken at various aspect ratios and angles, a Matlab script was created. Programming was used to combine the colors of the input photographs, and the output was created utilizing a variety of color combinations [4].

The authors of this study created, trained, and tested a model using an aerial image dataset. Classes from various car types and colors were included in the dataset. The authors designed the model using a deep CNN classifier-based methodology. The model compared the obtained aerial image with the expected class and produced an output of Yes or No, depending on whether a match was discovered or not [5].

The authors of this study present a method for identifying pedestrians and detecting their actions using aerial photos captured by a CNN model based on deep learning. To generate object proposals, authors used Single Shot Multi-box Detector (SSD) [6].

In order to recognize and identify objects, color is important. Authors have broken down difficult mathematical ideas and their significance for color usage in digital image processing in this book. The book's chapters go into great detail on various industry use cases as well as methods for utilizing color in DIP technologies [7].

The procedure for applying the Viola-Jones algorithm to face recognition has been described by the authors in this study. The algorithm was implanted and the outcomes were observed in real time by the authors through the integration of Matlab and Arduino software. Such use scenarios are often used in security surveillance [8].

The writers of this work have created a novel method for recognizing and locating the vegetables. This was created with the use case of huge warehouses and shopping centers that store veggies in mind. Making the mall checkout procedure more efficient and speedy was the goal. Several prominent properties, including weight, texture, and color, are first recorded when the photos are taken. After that, the vegetable is identified by analyzing these properties. The study's conclusions were thought to be promising [9].

The three approaches of colony counting with image processing techniques were tested and compared in this paper by the authors using Matlab and Photoshop. This high-end software's capabilities, effectiveness, and convenience of use were examined, and the findings demonstrated effective use cases for picture enhancing [10].

In this research, the authors present a computer vision approach for salient object recognition. Image extraction has been done using deep learning, a part of the AI field. When the authors evaluated their CNN-designed method on other datasets, they found that the deep learning-based model performed better than the other methods [11].

The authors of this paper used deep learning neural networks to determine the precise and right denomination of the currency notes by detecting features on both their front and back sides. With a 96.6% accuracy rate, the authors' Single Shot MultiBox Detector (SSD) model was able to determine the paper note's right denomination [12].

The YOLO model is an architecture based on neural networks that the authors have presented for object detection in this work. The entire detection workflow was established as a single network thanks to the architecture's clever design, which improves optimization performance and speeds up processing. Comparing the results to other widely used object detection techniques, including region-based convolutional neural networks, they were shown to be superior. At a rate of 155 frames per second, the suggested R-CNN based model could process images [13].

The application of deep CNN-based models to medical imaging difficulties has been examined by the authors in this work. Authors made observations after examining CNN designs with a wide range of layers and millions of parameters. Additional research was done to determine how the size of data sets and the spatial context of images affected the overall performance of the model. A sensitivity rate of 85% was achieved by the model when it was assessed for medical imaging cases [14].

The authors of this work present a single deep neural network-based method for image identification. The Single Shot MultiBox Detector technique operates on objects of different sizes and shapes, and all of the processing is done on a single network. When the model was assessed using a variety of public datasets, it was discovered that it performed faster and had a higher accuracy rate (76.9%) than other models [15].

Several deep learning techniques for pattern recognition have been examined, contrasted, and classed by the authors in this work. Researchers have found that knowing the benefits and particular applications of each model and tool can help them connect the problem statement to the appropriate model and instrument for potential implementation [16].

The authors of this work implemented the detection of human faces using the well-known Viola Jones approach. The Matlab software suite was used to build the programming and architecture of the face detection model. Researchers found that applying the technology to similar face detection scenarios was made easier by the documented step-by-step workflow execution. For different tuning parameters, the face detectors' performance was investigated [17].

A method for an open-ended Visual Question Answering (VQA) was suggested by the authors in this work. Offering a precise response in plain language to a query on the image and its context is one way to assist those who are visually impaired. The authors make available online a baseline, multiple choice questions and answers, and an image data collection [18].

The authors of this article examined CNN's behavior to determine how precisely to configure huge images. To evaluate the networks with continuously rising depth, a small-sized convolution filter architecture was created expressly for them. The outcomes were discovered to be encouraging, and other public datasets might make use of the findings [19].

The writers of this research have attempted to lower the error ratios while designing digital imaging systems for face recognition. The outcomes were found to be superior when compared to vendor tests for facial recognition [20].

The approach used to benchmark facial recognition system performance is discussed by the authors in this research. Face Recognition Vendor Test, or FRVT, is the test's official name. The data showed that, compared to females, males may be simpler to identify. Another intriguing finding was that young persons are harder to identify than elderly people [21].

The research of face detection in photographs was the authors' main area of interest in this paper. For the benefit of future researchers, a number of literature publications were examined, and face detection techniques were categorized and classed [22].

The authors of this work have developed a machine learning-based method for visual object recognition. The design achieves a greater rate of object detection and is very good at processing images quickly [23].

The authors of this study have presented a novel idea for pattern recognition. Their method, which is based on the bi-spectra approach, is well-known for managing Gaussian noise with improved immunity. The findings demonstrated an improvement in greater classification accuracy, even for cases with lower signal-to-noise ratios [24].

An analysis of research publications over the last five years also revealed that, like in the case of pattern identification in image recognition and detection, researchers have employed a variety of AI techniques, including ANN and ANFIS, to successfully identify the patterns of defects [25–27].

III. COMPARATIVE STUDY

The comparative analysis of object detection methods and observed outcomes is given in Table 1.

TABLE I. COMPARATIVE STUDY

Year	Ref. No	Comparative Analysis	
		Technique Used	***Results***
2019	[1]	Minimum distance classifier using a statistical distance measure	Higher order spectral invariants are a fast, parallel-implementable method for pattern recognition with strong resistance against additive Gaussian noise.
2018	[11]	Convolution Neural Network (CNN)	Findings across several datasets indicate that the deep learning-based model has outperformed alternative methods.
2018	[13]	Region-based Convolutional Neural Networks.	155 frames per second of image processing might be achieved using the suggested R-CNN based model.
2018	[12]	Using deep learning neural networks, one can identify features on both the front and reverse of money notes.	96.6% accuracy rate was attained.
2016	[15]	Single deep neural-network	Attained 76.9% accuracy utilizing a single-shot multibox detector
2016	[14]	Deep CNN based models in the area of medical imaging	Obtained a sensitivity rate of 85% using medical imaging instances.
2015	[17]	Viola Jones algorithm	The face detectors' performances for various tuning settings were examined.
2014	[19]	Convolution Neural Network (CNN)	The findings were deemed promising and might be applied to other publicly available datasets.
2003	[21]	Face recognition vendor test (FRVT).	Based on the findings, it appears that males may be simpler to recognize than females.
2001	[23]	Machine Learning Based Approach	Design is incredibly efficient at quickly processing images and producing results, like as a greater item detection rate.

IV. CONCLUSION

We provided a review research on object detection in this paper. This study presents an overview of object detection, followed by a summary based on a survey of twenty-four papers. We conducted a survey regarding several methods of object detection. This survey offers theoretical understanding of several object detecting techniques. Based on the results of this study, AI-based methods—in particular, CNN and Deep Neural Networks—have been shown to be successful and might be suggested for further object detection research.

V. REFERENCES

[1] AN. Jayanthi, C. Nareshkumar, S. Rajesh, "Fruit Quality Inspection System Using Digital Image Processing Technique", Iconic Research and Engineering Journals, 2, pp.260-263, 2019.

[2] S. Gong, C. Liu, Y. Ji, B. Zhong, Y. Li, H. Dong, "Visual Object Recognition. In Advanced Image and Video Processing Using MATLAB" pp. 351-387. Springer, Cham ., 2019.

[3] D. E. Holler, M. Behrmann, and J. C. Snow, "Real-world size coding of solid objects, but not 2-D or 3-D images, in visual agnosia patients with bilateral ventral lesions," Cortex, vol. 119, pp. 555–568, Oct. 2019.

[4] Z. Bhutto, MZ. Tunio, A. Hussain, J.Shah, I.Ali, MH. Shaikh, "Scaling of Color Fusion in Stitching Images," IJCSNS, 19(4), 2019.

[5] A. Soleimani, N. M. Nasrabadi, E. Griffith, J. Ralph, and S. Maskell, "Convolutional Neural Networks for Aerial Vehicle Detection and Recognition," NAECON 2018 - IEEE National Aerospace and Electronics Conference, Jul. 2018.

[6] A. Soleimani and N. M. Nasrabadi, "Convolutional Neural Networks for Aerial Multi-Label Pedestrian Detection," 2018 21st International Conference on Information Fusion (FUSION), Jul. 2018.

[7] A. M. Grigoryan and S. S. Agaian, "Quaternion and Octonion Color Image Processing with MATLAB," Apr. 2018.

[8] V.S., Pamulapati, Y.S. Rohan, V.S Kiran, S. Sandeep, and M.S. Rao, " Real-Time Face Tracking Using Matlab And Arduino". Electronics And Communication Engineering, Vasireddy Venkatadri Institute Of Technology, 2018

[9] S.Biswas, N.Ullah, M.Mazumder, "Machine Vision Based Vegetable Recognition", 2018.

[10] X. Chen, S. Li, Z. Zhang, and J. Gao, "The Colony Count Based on Image Processing Using Matlab and Photoshop," Proceedings of the 2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE 2018), 2018.

[11] F. Xiao, W. Deng, L. Peng, C. Cao, K. Hu, and X. Gao, "Multi-scale deep neural network for salient object detection," IET Image Processing, vol. 12, no. 11, pp. 2036–2041, Nov. 2018.

[12] Q. Zhang and W. Q. Yan, "Currency Detection and Recognition Based on Deep Learning," 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Nov. 2018.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016.

[14] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1285–1298, May 2016.

[15] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A.
C. Berg, "SSD: Single Shot MultiBox Detector," Lecture Notes in Computer Science, pp. 21–37, 2016.

[16] P. N. Druzhkov and V. D. Kustikova, "A survey of deep learning methods and software tools for image classification and object detection," Pattern Recognition and Image Analysis, vol. 26, no. 1, pp. 9–15, Jan. 2016.

[17] E. Alionte and C. Lazar, "A practical implementation of face detection by using Matlab cascade object detector," 2015 19th International Conference on System Theory, Control and Computing (ICSTCC), Oct. 2015.

[18] S. Antol, A. Agrawal, J. Lu, M. Mitchell, D. Batra, C. L. Zitnick, and
D. Parikh, "VQA: Visual Question Answering," 2015 IEEE International Conference on Computer Vision (ICCV), Dec. 2015.

[19] T. Pfister, K. Simonyan, J. Charles, and A. Zisserman, "Deep Convolutional Neural Networks for Efficient Pose Estimation in Gesture Videos," Lecture Notes in Computer Science, pp. 538–552, 2015.

[20] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the Face Recognition Grand Challenge," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).

[21] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi,and M. Bone, "Face recognition vendor test 2002 :," 2003.

[22] Ming-Hsuan Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34–58, 2002.

[23] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.

[24] V. Chandran and S. L. Elgar, "Pattern Recognition Using Invariants Defined From Higher Order Spectra- One Dimensional Inputs," IEEE Transactions on Signal Processing, vol. 41, no. 1, p. 205, Jan 1993.

[25] V. Vashisht, M. Lal, and G. S. Sureshchandar, "A Framework for Software Defect Prediction Using Neural Networks," Journal of Software Engineering and Applications, vol. 08, no. 08, pp. 384–394, 2015.

[26] V. Vashisht, M. Lal, and G. Sureshchandar, "Defect Prediction Framework Using Neural Networks for Software Enhancement Projects," British Journal of Mathematics & Computer Science, vol. 16, no. 5, pp. 1–12, Jan 2016.

[27] V. Vashisht, M. Lal, and G. Sureshchandar, "Defect Prediction Framework Using Adaptive Neuro-Fuzzy Inference System (ANFIS) for Software Enhancement Projects," British Journal of Mathematics & Computer Science, vol. 19, no. 2, pp. 1–12, Jan 2016.