The uses of Computer Vision in Artificial Intelligence to extract information from images

Ali N. Yousif

University of Information Technology and Communications, Baghdad, Iraq
alinafea@uoitc.edu.iq

ABSTRACT

Machine learning is made possible by the accessibility and availability of sizable data gathered from the sensors, internet, and Modern technology devices in this field. The main idea of machine learning explains the possibility of a computer improving itself over time. Pattern recognition, computer vision, and other learning processes can be aided by organizing conferences, seminars, and group discussions. Experimentation and real-world applicability, this research in computer visions and the concept of machine learning investigates and analyses, assesses, and forecasts the potential of computer vision machine learning applications. The research has discovered that monitored, unmonitored, and machine learning algorithms with semi- monitored are used in computer vision. Support vector machines, neural networks, and k-means clustering are the most widely used methods.

Keywords: Machine learning, computer vision, deep learning, and artificial intelligence

Introduction:

The goal of machine learning and using computer vision simultaneously is to enable computers to sense data, comprehend data, and respond based on previous and present findings, researches are still constantly evolving [1]. For the (IoT) Internet of Things, Industrial Internet of Things, and brain-computer interfaces, computer vision is crucial. These techniques are used to recognize and track complex human behaviors in multimedia streams. There are numerous tried-and-true techniques for analysis and prediction, including monitored, unmonitored, and semi monitored learning. These techniques use a variety of machine learning algorithms, such as KNN and support vector machines.

The main components of machine learning solutions are data collection, model training, and prediction using the trained model Private businesses offer models and services for image classification, text analysis, and speech recognition. By application programming interfaces (API), these models may be accessed through services like Microsoft Azure Cognitive Services, IBM Watson, Polly, Amazon Recognition, and Lex. One of the most important aspects of daily life is finding and examining objects. Application areas for object detection include preventing road accidents, identifying emotions from human postures, and recognizing facial expressions. Created an automated method that uses orientation to extract the data from human faces in photos and videos [2]. Software libraries utilized in computer vision and object recognition is Tensor Flow and Open Pose. Long short-term memories, gated recurrent units, recurrent neural networks, and convolutional neural networks, and Bayesian networks are all used in traffic detection models. Sensors in intelligent settings collect data that are then analyzed and forecasted[3][4]. One of the tasks that the convolution neural network (CNN) successfully completes for successful object detection is feature extraction. With a huge collection of face

photos, through monitored learning, a deep convolutional neural network is able to recognize faces [5]. The data annotation/labeling is the only issue with machine learning applications and computer vision [6]. Currently, "cloud machine learning" or "machine learning as a service" is how the machine learning algorithms are executing on the cloud. Additionally, businesses like Google, Microsoft, and Amazon provide machine learning such as the concepts of cloud service. The goal of this paper project is to look into and critically assess machine learning and techniques of computer vision.

The database that was searched includes scientific browsers, and sophisticated search ways were applied to the keywords "machine learning" and "computer vision" Deep learning and Artificial intelligence 258 papers, including patents and citations, were found in the initial search. The number was reduced to 175 articles after the text of the publications was examined and the citations were disregarded

Background study:

Recent research has concentrated on a number of significant subjects, including computer vision and machine learning. Picture and pattern mappings are used by the computer vision system to find answers [7]. It sees an image as a collection of pixels. The monitoring, inspection, and surveillance operations are automated by computer vision [5]. A result of dealing with machine learning and developing the techniques of computer vision is the automated annotation and analysis of videos. Classification, object identification, and instance segmentation are displayed in Figure 1. Figure 2 displays object recognition in images using the Faster-RCNN-Inception-V2 model with the Tensor flow in the Anaconda environment.

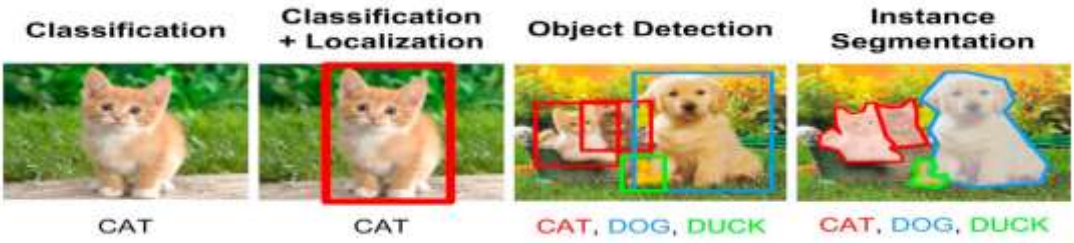

Figure 1: automated annotation and analysis of videos

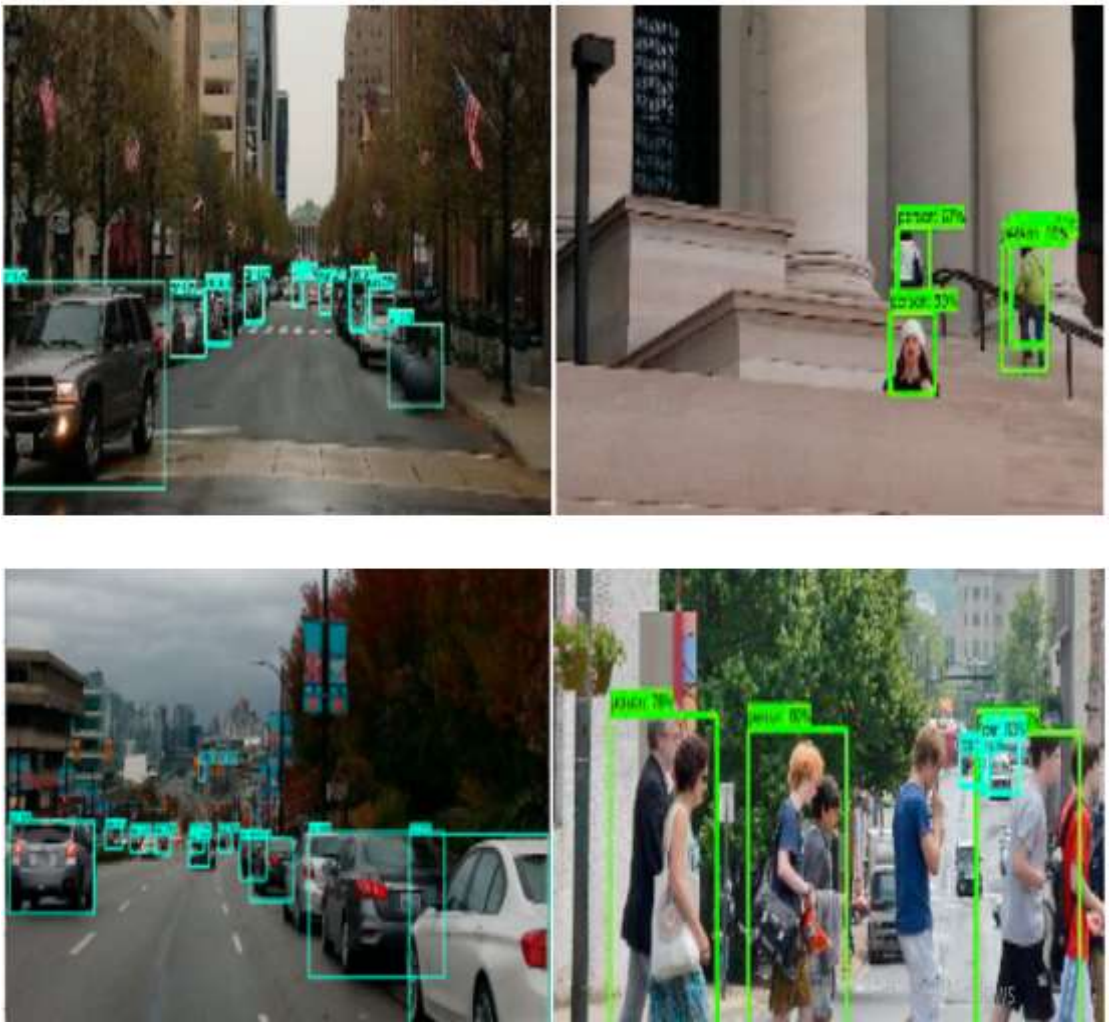

Figure. 2: How to detect autos and people in photos Faster-RCNN-Inception-V2

In this domain of machine learning, there are three types of learning: monitored, unmonitored, and semi- monitored. Monitored learning was used to label the training data. The process of data labeling is costly, time-consuming, and labor-intensive. However, with semi- monitored learning, some of the data is labeled and some is not. When learning from unlabeled data, the Bayesian network classifiers have the advantage. However, the actual the category of unmonitored learning in which patterns develop based on clustering includes problems.

The paradigms include support vector machines, neural networks, and probabilistic graphical models. Support vector machines (SVMs) are a popular monitored machine learning technique for categorization [8]. Layered networks of connected processing nodes make up a neural network [9]. A class of neural networks called convolutional neural networks (CNNs) is utilized for image recognition and categorization. It has neurons that are wide, tall, and deep. Due to widely available datasets, GPUs, and regularization approaches, CNN has grown in popularity recently. For image processing and analysis, the Open library may be integrated with platforms like Eclipse and Visual Studio in Windows, iOS, and Linux, as well as programming languages including Android,.NET, Java, and iOS.

Applications of the study:

The study investigated a number of computer vision uses of machine learning. Pattern matching, Feature extraction from surface reconstruction, representation, and modeling for biological sciences are a few examples of how to segment data and refine visual models. Machine learning uses in addition to computer

vision to understand data in photos that recognize cars and pedestrians [10], automatically classify railway ties using photographs [8], and more. Separating mango varieties based on the analysis of far recognition data for geographic information systems, the extraction of linguistic information from document pictures [11], and size attributes. Likewise, other applications include handwriting recognition, machine vision, gesture and face recognition, and face and gesture recognition Numbers [12], cutting-edge driving assistance technologies [13], behavioral research, and assessment of the full-body size of a person's bicycle kinematics and pose estimation [14], also, examined how are used in fields of medicine such microscopy, magnetic resonance imaging, endoscopy, thermography, angiography, and nuclear medicine [15]

Results:

Contrary to textual information, which is constantly changing in the online world, it takes a lot of labor to be able to categorize and maintain them according to their unique qualities. Computer intrusions with high model-based vision skills and learnability are needed to index and preserve graphical data. This paper highlights various areas, such as engineering, research, and technology-related costs, efforts. The probabilistic models use labeling and pattern recognition to forecast human behavior. In professional sports, player and team performance is assessed using machine learning in general and benefit from the techniques of computer vision. Preventive maintenance has also been employed in other sectors. The efficacy and efficiency of enterprises are greatly impacted by replacing equipment and tools before they break down. Key data sources include the network of public cameras and smart gadgets with sensors. Using these data and the algorithms of machine learning, it is possible to forecast and track city traffic. The growing fields of machine learning uses with the cooperation of computer vision research are depicted in Figure 3. Traffic management and biological sciences are the two areas of study in this area that are growing at the fastest rates (13 percent)

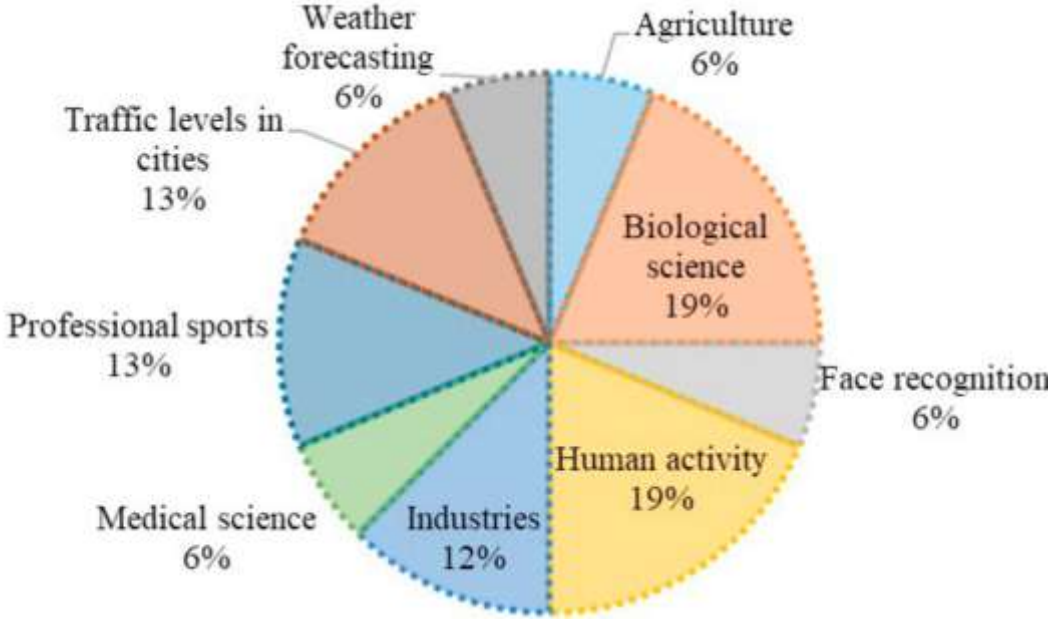

Figure 3: research fields

In the realm of machine learning, more complex approaches to image interpretation have replaced conventional methods of pattern recognition and image processing. It has a strong probability of influencing future developments in computer vision systems. Nevertheless, despite machine learning's ability to do predictive analysis on assembled data, computer vision can comprehend and extract information from audio and video and be employed independently of it. On the basis of the photographs alone, it is challenging to distinguish between fire and explosion. In the past three years, research has been conducted

in a variety of fields, including agriculture, biology, medicine, psychology, industrial maintenance, sports analysis and forecasting, and traffic management and control in cities. Brain computer interfaces have received scant attention in computer vision domains in general and machine learning topics research. Inputs for machine learning techniques within computer vision can also be either direct inputs (such as pixels, voxels, or 3D points) or vectors (measures of shape, distributions of edges and colors, and measurements and distributions of texture). In vision applications, vectors are typically employed to express the properties. Researchers studying pattern recognition employ graphs rather than machine learning, which prefers first order logic formalism, to represent the structure of objects and scenes [16].

Conclusion

The development of novel methods, procedures, and technologies is a result of both commercial and university computer vision research. Many problems with feature extractions and feature selection have been solved using machine learning and working with visual processing in computers. The applications of machine learning differ in their results according to the domain, and computer vision also differs as a result of the field of work. The application of machine learning in computers is analyzed, classified, and discussed in this paper. The study has identified the effective use of computer applications for machine learning vision for the biological sciences, food security, expression reading, species taxonomy, sports, and forecasting weather traffic flow monitoring, industrial predictive maintenance. The growing fields include biology, the interpretation of human behavior, traffic control, and professional sports. The most common computer vision uses of machine learning are for object recognition, classification, and prediction.

References:

[1] N. Kerle, M. Gerke, and S. Lefèvre, "GEOBIA 2016: Advances in object-based image analysis-linking with computer vision and machine learning," *Remote Sens.*, vol.11, no.10, pp. 2016–2018, 2019, doi :10.3390/rs11101181.

[2] N. Sebe, M. S. Lew, Y. Sun, I. Cohen, T. Gevers, and T. S. Huang, "Authentic facial expression analysis," *Image Vis. Comput.*, vol. 25, no. 12, pp. 1856–1863, 2007, doi: 10.1016/j.imavis.2005.12.021.

[3] I. M. Kaloskampis, "Recognition of Complex Human Activities in Multimedia Streams using Machine Learning and Computer Vision," 2013, [Online]. Available: http://orca.cf.ac.uk/id/eprint/59377.

[4] A. Al-Badi, A. Tarhini, and A. I. Khan, "Exploring big data governance frameworks," *Procedia Comput. Sci.*, vol. 141, pp. 271–277, 2018, doi: 10.1016/j.procs.2018.10.181.

[5] R. Bond *et al.*, "Democratisation of Usable Machine Learning in Computer Vision," pp. 1–4, 2019, [Online]. Available: http://arxiv.org/abs/1902.06804.

[6] A. Fathi, "Recent Hot Machine Learning Hammers used in Computer Vision," *Cc.Gatech.Edu*, no. Mil, pp. 1–6, 2012, [Online]. Available: http://www.cc.gatech.edu/~afathi3/publication/RecentHammers.pdf.

[7] F. Esposito and D. Malerba, "Editorial: Machine learning in computer vision," *Appl. Artif. Intell.*, vol. 15, no. 8, pp. 693–705, 2001, doi: 10.1080/088395101317018546.

[8] A. Kulkarni, D. Parikh, and A. L. Abbott, "Classification of Faults in Railway Ties using Computer Vision and Machine Learning," 2017.

[9] A. Drory, "Computer Vision and Machine Learning for Biomechanics Applications Human Detection, Pose and Shape Estimation and Tracking in Unconstrained Environment From Uncalibrated Images, Videos and Depth," no. September, 2017.

[10] C. Papageorgiou, "A Trainable System for Object Detection in Images and Video Sequences," *Int. J. Comput. Vis.*, vol. 38, no. 1685, pp. 15–33, 2000.

[11] F. Esposito, D. Malerba, and F. A. Lisi,

"Machine learning for intelligent processing of printed documents," *J. Intell. Inf. Syst.*, vol. 14, no. 2–3, pp. 175–198, 2000, doi: 10.1023/A:1008735902918.

[12] T. Dash and T. Nayak, "English Character Recognition using Artificial Neural Network," *CoRR*, vol. abs/1306.4, pp. 7–9, 2013.

[13] B. Huval *et al.*, "An Empirical Evaluation of Deep Learning on Highway Driving," pp. 1–7, 2015, [Online]. Available: http://arxiv.org/abs/1504.01716.

[14] S. Hara, S. Kobayashi, and M. Abe, "Sound collection systems using a crowdsourcing approach to construct sound map based on subjective evaluation," *2016 IEEE Int. Conf. Multimed. Expo Work . ICMEW 2016* , 2016 , doi:10.1109/ICMEW.2016.7574694.

[15] J. Antonio and P.-´ Alava, "Computer Vision and Medical Image Processing: A brief survey of application areas," pp. 152–159, 2015.

[16] J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl, and V. Svetnik, "Deep neural nets as a method for quantitative structure-activity relationships," *J. Chem. Inf. Model.*, vol.55, no.2, pp.263–274, 2015, doi:10.1021/ci500747n.