Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 1444–1451

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Machine Learning in Computer Vision

Asharul Islam Khan [a]*, Salim Al-Habsi[b]

[a] *Remote Sensing and GIS Research Center*
Sultan Qaboos University, P.O.Box 33 AlKhodh, Muscat, P.C.123, Oman
[b] *Gulf College, P.O.Box 885, Al Khuwair, Al Mabaila, Muscat, P.C. 123, Oman*

Abstract

During last few years the computer applications have gone dramatic transformation from simple data processing to machine learning, thanks to the availability and accessibility of huge volume of data collected through sensors and internet. The idea of machine learning demonstrates and propagates the facts that computer has the ability to improve itself with the passage of time. The western countries have shown great interest on the topic of machine learning, computer vision, and pattern recognition via organizing conferences, workshops, collective discussion, experimentation, and real life implementation. This study on machine learning and computer vision explores and analytically evaluates the machine learning applications in computer vision and predicts future prospects. The study has found that the machine learning strategies in computer vision are supervised, un-supervised, and semi-supervised. The commonly used algorithms are neural networks, k-means clustering, and support vector machine. The most recent applications of machine learning in computer vision are object detection, object classification, and extraction of relevant information from images, graphic documents, and videos. Additionally, Tensor flow, Faster-RCNN-Inception-V2 model, and Anaconda software development environment used to identify cars and persons in images.

© 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019).

Keywords: machine learning; image processing; object detection; computer vision; artificial intelligence; image classification; neural network; support vector machine.

* Corresponding author. Tel.: +968-95687668.
E-mail address: ashar.367@gmail.com

1877-0509 © 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019).
10.1016/j.procs.2020.03.355

1. Introduction

The machine learning and computer vision hopes to bring into the computers the human capabilities for data sensing, data understanding, and action taking based on the past and present outcomes. The machine learning and computer vision research is still evolving [1]. Computer vision is an essential part of Internet of Things, Industrial Internet of Things, and brain human interfaces. The complex human activities are recognized and monitored in multimedia streams using machine learning and computer vison. There are numbers of well-established methods for prediction and analysis such as supervised learning, un-supervised learning, and semi supervised learning. These methods uses the machine learning algorithms such as support vector machine, KNN etc.

The machine learning solutions revolves around data gathering, training a model, and use the trained model to make predictions. There are models and services provided by private companies for speech recognition, text analysis, and image classification. One can use their models through application programming interfaces (API). For instance, Amazon Recognition, Polly, Lex, Microsoft Azure Cognitive Services, IBM Watson. Object detection and analysis is an important part of everyday life. Object detection has applications in avoiding traffic collisions, facial expression recognition and emotional recognition based on human postures. In [2], developed an automated system to detect the information contained in human faces from images and video with the help of orientations**.** The TensorFlow and OpenPose are the software library used in object detection and computer vision. The traffic detection models uses convolutional neural network, recurrent neural network (RNN), long short-term memory (LSTM), gated recurrent unit (GRU), and Bayesian networks. In an intelligent environments sensors capture the data which is later used for analysis and forecasting [3, 4]. The feature extraction is one of the task convolution neural network (CNN) accomplishes without information loss for successful object detection [5]. The supervised learning of a deep convolutional neural network recognizes faces with a large set of face images [6]. The only challenge in computer vision and machine learning application is the data annotation/ labelling [7]. The machine learning algorithms are now running on cloud as a "machine learning as a service", "cloud machine learning" [6]. Moreover, companies, such as Amazon, Microsoft, and Google, have machine learning as a cloud service.

The objective of this research study is to investigate and analytically evaluate applications of machine learning in computer vision. The database searched includes Google scholar applying advanced search techniques with respect to keywords- "machine learning", "computer vision", "Deep learning", and "Artificial intelligence". The initial search resulted in 258 articles, which included both patent and citation. After examining articles contents and excluding the citations, the number came down to 175 articles. Finally, 20 articles formed the core part of this research study. There are five sections. Section 2 corresponds to background study. Section 3 clusters the existing the machine learning applications in groups. Section 4 presents results and discussions. The last section concludes with comments and future work.

2. Background study

The computer vision and machine learning are two important areas of recent research. The computer vision computer uses the image and pattern mappings in order to find solutions [8]. It considers an image as an array of pixels. The computer vision automates the monitoring, inspection, and surveillance tasks [6]. Machine learning is the subset of artificial intelligence. The automatic analysis/annotation of videos is the outcome of computer vision and machine learning. Figure 1 shows the classification, object detection, and instance segmentation. Figure 2 shows the object detection in images using Tensor flow and Faster-RCNN-Inception-V2 model in Anaconda environment.

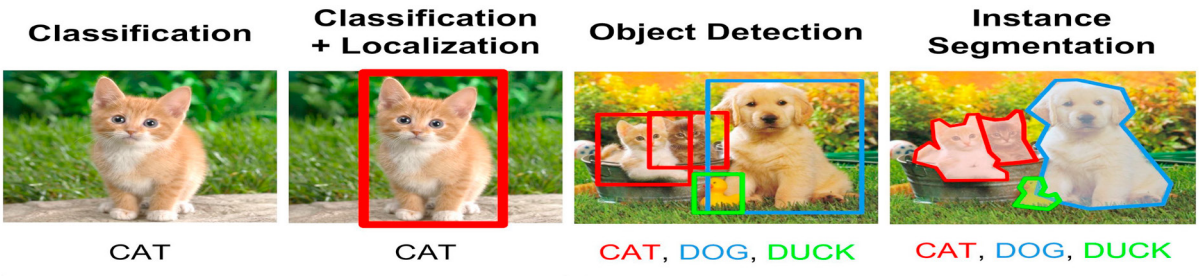

Fig 1. Classification, object detection, and instance segmentation [9]

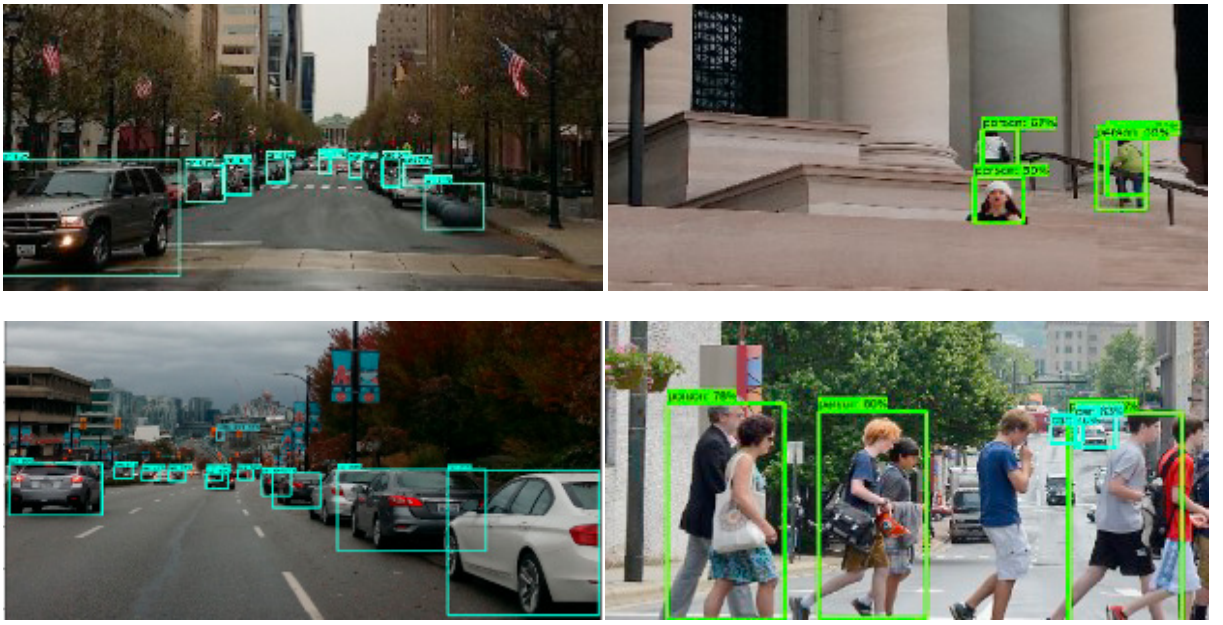

Fig 2. Detecting cars and persons in images applying Deep learning and Faster-RCNN-Inception-V2 model

There are three approaches to machine learning and computer vision: supervised, unsupervised, and semi-supervised learning. The supervised learning has labelled training data. The labelling of data is expensive, time consuming and requires expertise. On the other hand the semi-supervised learning has some of the data labelled and others not. The Bayesian network classifiers has the edge for learning with unlabelled data. Nevertheless, the real world problems falls in the unsupervised learning category where patterns evolves based on clustering.

The machine learning paradigms for computer vision are support vector machines, neural networks, and probabilistic graphical models. Support vector machines (SVMs) is a subdomain of supervised machine learning methods and popular in classification [10]. Neural network consists of layered networks of interconnected processing nodes [11]. Convolutional neural networks (CNNs) is a category of neural networks used in image recognition and classification. It has neurons with dimensions: width, height and depth [10]. CNN has gained popularity recent times due to largely accessible datasets, GPUs, and regularization techniques [10]. OpenCV is a library, which can be integrated with programming languages such as Android, .NET, Java, iOS on platforms such as Eclipse and Visual Studio in Windows, iOS, and Linux for image processing and analysis. It is used in image processing, video analysis, object detection, and machine learning. Figure 3 shows the object detection process in the machine learning and computer vision environment.

Fig 3. Image processing and object detection process

3. Machine learning in Computer vision

The study explored numerous applications of machine learning in computer vision. For example, segmentation, feature extraction, refining visual models, pattern matching, shape representation, surface reconstruction, and

modelling for biological sciences. Machine learning in computer vision is used in interpreting data contained in car and pedestrian detection images [12], automatic classification of faults in railway ties using images [10], interpretation of remote sensing data for geographical information systems [13], differentiating mango verities based on size attributes [14], extraction of graphical and textual information from document images [15]. Similarly, other applications includes gesture and face recognition [16], machine vision [17], recognizing handwritten characters and digits [18], advanced driver assistance systems [19], behavioural studies [20], and estimation of human full-body kinematics for a cyclist and pose estimation [11]. Detecting Curb Ramps in Google Street View such as automatically identification and examination of curb ramps in images [21]. In [22], studied use of computer vision and machine learning in medical science such as cardiovascular imaging, retinal blood vessels, nuclear medicine, endoscopy, thermography, angiography, magnetic resonance, ultrasound and microscopy. Machine learning and computer vision has innovative applications in engineering, medicines, agriculture, astronomy, sports, education etc. These applications are categories and clustered in Table 1.

Table 1. Machine learning and computer vision research

Researchers	Demonstrated application area	Description
[23, 24]	Food security, agricultural production, flood prediction, and oil palm tree counting.	The agricultural fields and land cover is mapped after processing the satellite images. For example, Mapping Sub-Saharan African Agriculture, the satellite images are processed and classified using machine learning algorithm such as random forest [23]. In [24], proposed a method to detect, differentiate, and count the oil palm tree using machine learning and computer vision. The method worked 96 % accurately.
[25]	Rainfall, flood, wind, temperature, humidity, and front detection.	The computing power is used for fast and accurate weather forecasting. The computer vision and machine learning is used in the front detection (meteorological phenomena where two distinctly different air masses meet and interact) for weather forecasting [25]
[26-28]	Occupancy detection, traffic detection, tracking, classification and counting.	The detection of traffic flow on a road and categorization viz., cars, bicycles, and pedestrians [26]. For example, the traffic in Montreal is predicted using computer vision and machine learning techniques [27]. In [28], developed an end-to-end system to detect, track, count, and manage traffic (pedestrians and bicyclists) in Los Angles using machine learning and computer vision.
[29, 30]	Haemorrhoid detection, Bleeding detection, endoscopic image enhancement, and clinical decision support	In [29] used computer vision and machine learning for breast cancer diagnosis. The computer vision and machine learning are employed in gastrointestinal (GI) endoscopy [30].
[2, 31, 32]	Human behaviours, face-to-face conversations, emotion recognition, and phone conversations.	The complex human behaviours' are modelled using Bayesian networks. The semantic events from audio-visual data with spatial-temporal support detected by fusing the information extracted from multiple modalities [2]. In [31], proposed an approach to detect and classify human behaviours' as confident/ not confident based on human posture using machine learning and computer vision. In [32], proposed Gesture Learning Module Architecture (GeLMA), for hand gesture recognition in real time. The architecture proved successful with 99 % accuracy.
[33-35]	Classification of biological fluorescence images of synapses, protein localization,	The k-nearest neighbour (kNN) classification for protein localization, Naïve Bayes Classifiers for phylogenetic reconstructions [35]. The traditional method to evaluate crop

	phenotyping, and phylogenetic reconstructions	biotic and abiotic stresses are time taking and requires a lot of efforts [33, 34].
[36, 37]	Performance evaluation, scoreboard updating, and predicting the game outcomes.	In [36], automated the cricket scorecard with computer vision and machine learning. This is useful for updating the cricket scoreboard based on the umpire gesture in the field. Likewise, in-Play Tennis Analysis, analysing the players performance and predicting the future outcomes [37].
[38]	State of machine tools, repairing, predictive maintenance.	In [38], identified state of tools in milling process. The tools used in milling process their states, as a wear level is analysed and classified using computer vision and machine learning.

4. Results and discussion

In the world of internet, tons of graphical information and images moves but unlike the textual data the capability to classify and store them according to the special characteristics is a labour intensive task. The indexing and storing of graphical data requires computer interventions with advanced model based vision capabilities and learnability.

This study highlights the research in machine learning and computer vision in various domains. The machine learning and computer vision techniques has reduced the cost, effort, and time in engineering, science, and technology. An automated system based on machine learning and computer vision detects the human emotions (likes and dislikes, confidence levels). The probabilistic models predicts human activities through labelling and pattern recognition. The machine learning and computer vision in professional sports measures and analyses the performances of team and individual players. Moreover, it has been in use in industries for predictive maintenance. The machine and tools replacements in industries on time before failures have significant impact on the effectiveness and efficiency of the manufacturing units. The public camera and smart devices with sensors are huge source of data. The computer vision and machine learning techniques when applied on these data helps in prediction and monitoring the traffics in cities. Figure 4 shows the evolving research areas in machine learning and computer vision. This study has found advancing areas of research in this field as- the biological science (19%) and human activity (19%) followed by traffic management (13%) and professional sports (13%).

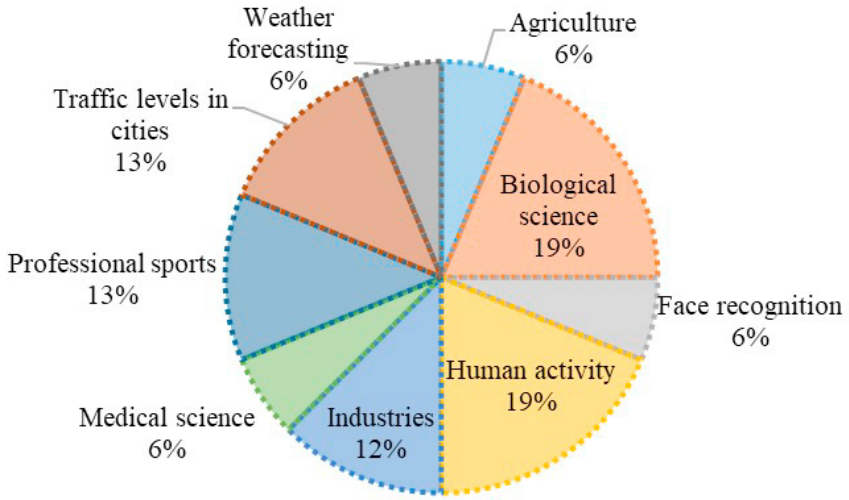

Fig 4. Machine learning and computer vision research areas

Machine learning field has evolved from the traditional methods of pattern recognition and image processing to advanced techniques of image understanding. It has a strong potential to contribute to the changing dynamic of

computer vision system. Although, the computer vision interpret and extract information from audio and video and can work independently of machine learning, however machine learning adds the predictive feature of already processed data. Based on the pure visuals it is difficult to differentiate between fire and explosion. Figure 5 shows the undergoing research in machine learning and computer vision with respect to time. The majority of the research in the area started after 2015. During last three years, the research has been on, agriculture, biological and medical science, human activity interpretation, predictive maintenance in industries, analysis and prediction of sports, traffic management and control in the cities. The research works of machine learning and computer vision on brain computer interfaces are scant.

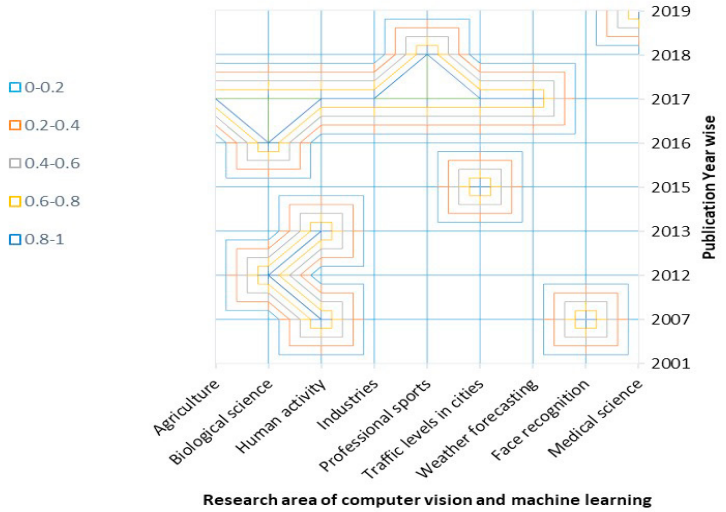

Fig 5. Distribution of machine learning and computer vision research with respect to time (years)

Moreover, the inputs to machine learning in computer vision is either of the form as a direct input (Pixels / Voxels / 3D Points) or of vectors (shape measures, edge distributions, colour distributions, texture measures / distributions). Vectors represents the features in many applications for vision. The researchers in pattern recognition captures the structure of objects and scenes using graphs as opposed to machine learning where first order logic formalism is preferred [8]. Figure 6 shows the frequent used themes of machine learning in computer vision. Machine learning in computer vision works for object classification, object detection, instance recognition, sequence recognition/ classification.

Fig 6. Frequently used themes of machine learning in computer vision

The learning algorithms helps in creating object and image detection systems based on the instances and experiences. The machine learning algorithms renders enormous capabilities for integration and synthesis of vision algorithms and models. Object detection and tracking is still open challenge in computer vision, even though alignment with machine learning algorithms and open source libraries have tendered exciting results. Moreover, the machine learning output quality depends on the predictive accuracy, recall, and precision. The data collected from the environment dictates the strategies that a learning system would use so to improve the performance.

5. Conclusion

The commercial and academic research on computer vision is growing in the form of new techniques, process, models, and algorithms. Machine learning has been able to address many issues of feature extractions and processing in computer vision. The machine learning and computer vision synthesis has helped in understanding complex problems. The machine learning applications in computer vision have varied outputs depending upon the domain. This study includes the analysis, classification, and discussion on the use of machine learning in computer vision. The research has identified the successful implementation of machine learning applications in computer vision for weather forecasting, biological science, expression reading, food security, species classification, sports, monitoring the traffic flows, and predictive maintenance in industries. The biological science, human activity interpretation, traffics management, and professional sports are the emerging areas. The object detection, classification, and prediction are the most frequent use of machine learning in computer vision. The future work would assess accuracy of the machine learning algorithms in computer vision.

References

[1] Kerle, Norman, Markus Gerke, and Sébastien Lefèvre, *GEOBIA 2016: Advances in Object-Based Image Analysis—Linking with Computer Vision and Machine Learning*. 2019.

[2] Sebe, Nicu, Michael S Lew, Yafei Sun, Ira Cohen, Theo Gevers, and Thomas S Huang, *Authentic facial expression analysis.* Image and Vision Computing, 2007. **25**(12): p. 1856-1863.

[3] Kaloskampis, Ioannis, *Recognition of complex human activities in multimedia streams using machine learning and computer vision*. 2013, Cardiff University.

[4] Al-Badi, Ali, Ali Tarhini, and Asharul Islam Khan. *Exploring Big Data Governance Frameworks*. in *9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2018)*. 2018. Leuven, Belgium: Procedia Computer Science.

[5] Zhang, Fan, Wei Li, Yifan Zhang, and Zhiyong Feng, *Data Driven Feature Selection for Machine Learning Algorithms in Computer Vision.* IEEE Internet of Things Journal, 2018. **5**(6): p. 4262-4272.

[6] Bond, Raymond, Ansgar Koene, Alan Dix, Jennifer Boger, Maurice D Mulvenna, Mykola Galushka, Bethany Waterhouse Bradley, Fiona Browne, Hui Wang, and Alexander Wong, *Democratisation of Usable Machine Learning in Computer Vision.* arXiv preprint arXiv:1902.06804, 2019.

[7] Fathi, Alireza, *Recent Hot Machine Learning Hammers used in Computer Vision.* 2012.

[8] Esposito, Floriana and Donato Malerba, *Machine learning in computer vision.* Applied Artificial Intelligence, 2001. **15**(8): p. 693-705.

[9] Ouaknine, A. *Review of Deep Learning Algorithms for Object Detection*. 2018 [cited 2019 11 June]; Available from: https://medium.com/zylapp/review-of-deep-learning-algorithms-for-object-detection-c1f3d437b852.

[10] Kulkarni, Amruta Kiran, *Classification of Faults in Railway Ties using Computer Vision and Machine Learning*. 2017, Virginia Tech.

[11] Drory, Ami, *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.* 2017.

[12] Papageorgiou, Constantine and Tomaso Poggio, *A trainable system for object detection.* International journal of computer vision, 2000. **38**(1): p. 15-33.

[13] Jung, Franck, *Detecting building changes from multitemporal aerial stereopairs.* ISPRS Journal of Photogrammetry and Remote Sensing, 2004. **58**(3-4): p. 187-201.

[14] Patel, Krishna Kumar, Abhijit Kar, and MA Khan, *Development and an Application of Computer Vision System for Nondestructive Physical Characterization of Mangoes.* Agricultural Research: p. 1-16.

[15] Esposito, Floriana, Donato Malerba, and Francesca A Lisi, *Machine learning for intelligent processing of printed documents.* Journal of Intelligent Information Systems, 2000. **14**(2-3): p. 175-198.

[16] White, David, James D Dunn, Alexandra C Schmid, and Richard I Kemp, *Error rates in users of automatic face recognition software.* PloS one, 2015. **10**(10): p. e0139827.

[17] Steger, Carsten, Markus Ulrich, and Christian Wiedemann, *Machine vision algorithms and applications*. 2018: John Wiley & Sons.

[18] Dash, Tirtharaj and Tanistha Nayak, *English Character Recognition using Artificial Neural Network.* arXiv preprint arXiv:1306.4621, 2013.

[19] Huval, Brody, Tao Wang, Sameep Tandon, Jeff Kiske, Will Song, Joel Pazhayampallil, Mykhaylo Andriluka, Pranav Rajpurkar, Toki Migimatsu, and Royce Cheng-Yue, *An empirical evaluation of deep learning on highway driving.* arXiv preprint arXiv:1504.01716, 2015.

[20] Garg, Shree, Ankush K Singh, Anil K Sarje, and Sateesh K Peddoju. *Behaviour analysis of machine learning algorithms for detecting P2P botnets*. in *2013 15th International Conference on Advanced Computing Technologies (ICACT)*. 2013: IEEE.

[21] Hara, Sunao, Shota Kobayashi, and Masanobu Abe. *Sound collection systems using a crowdsourcing approach to construct sound map based on subjective evaluation*. in *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. 2016: IEEE.
[22] Párraga Álava, Jorge Antonio. *Computer Vision and Medical Image Processing: A brief survey of application areas*. in *Argentine Symposium on Artificial Intelligence (ASAI 2015)-JAIIO 44 (Rosario, 2015)*. 2015.
[23] Debats, Stephanie Renee, *Mapping Sub-Saharan African Agriculture in High-Resolution Satellite Imagery with Computer Vision & Machine Learning*. 2017, Princeton University.
[24] Mubin, Nurulain Abd, Eiswary Nadarajoo, Helmi Zulhaidi Mohd Shafri, and Alireza Hamedianfar, *Young and mature oil palm tree detection and counting using convolutional neural network deep learning method.* International Journal of Remote Sensing, 2019: p. 14.
[25] Karlsen, Simen Skaret, *Automated Front Detection-Using computer vision and machine learning to explore a new direction in automated weather forecasting*. 2017, The University of Bergen.
[26] Mitra, Shounak, *Applications of Machine Learning and Computer Vision for Smart Infrastructure Management in Civil Engineering.* 2017.
[27] Forbes, Ian and Amber Gamrat. *Estimating Traffic Levels in Montreal using Computer Vision and Machine Learning Techniques*. 2015 [cited 2019 10 may]; Available from: http://rl.cs.mcgill.ca/comp598/fall2014/comp598_submission_95.pdf.
[28] Wang, Haiyan, Hunter Owens, Janna Smith, William Chernicoff, Mehran Mazari, and Mohammad Pourhomayoun. *An End-to-End Traffic Vision and Counting System Using Computer Vision and Machine Learning: The Challenges in Real-Time Processing*. in *The Third International Conference on Advances in Signal, Image and Video Processing*. 2018. Nice, France: IARIA.
[29] Yunchao, Gu and Yang Jiayao. *Application of Computer Vision and Deep Learning in Breast Cancer Assisted Diagnosis*. in *Proceedings of the 3rd International Conference on Machine Learning and Soft Computing*. 2019: ACM.
[30] Vemuri, Anant S, *Survey of Computer Vision and Machine Learning in Gastrointestinal Endoscopy.* arXiv preprint arXiv:1904.13307, 2019.
[31] Sacchetti, Rui, iago Teixeira, Bruno Barbosa, António J. R. Neves, Sandra C. Soares, and Isabel D. Dimas. *Human Body Posture Detection in Context: The Case of Teaching and Learning Environments*. in *The Third International Conference on Advances in Signal, Image and Video Processing*. 2018. Nice, France: IARIA.
[32] Trigueiros, Paulo José de Albuquerque C., *Hand Gesture Recognition System based in Computer Vision and Machine Learning : Applications on Human-Machine Interaction*, in *Tese de Doutoramento Electronic and Computer Engineering*. 2013, Universidade do Minho Escola de Engenharia.
[33] Zhang, Jiaoping, Hsiang Sing Naik, Teshale Assefa, Soumik Sarkar, RV Chowda Reddy, Arti Singh, Baskar Ganapathysubramanian, and Asheesh K Singh, *Computer vision and machine learning for robust phenotyping in genome-wide studies.* Scientific reports, 2017. **7**: p. 44048.
[34] Navarro, Pedro, Fernando Pérez, Julia Weiss, and Marcos Egea-Cortines, *Machine learning and computer vision system for phenotype data acquisition and analysis in plants.* Sensors, 2016. **16**(5): p. 641.
[35] Cooper, Roshni, *Accelerating neuronal genetic research in C. elegans with computer vision & machine learning.* CS231A Final Project Report, 2012.
[36] Shahjalal, Md Asif, Zubaer Ahmad, Rushrukh Rayan, and Lamia Alam. *An approach to automate the scorecard in cricket with computer vision and machine learning*. in *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)*. 2017: IEEE.
[37] Vinyes Mora, Silvia, *Computer vision and machine learning for in-play tennis analysis: framework, algorithms and implementation*. 2018, Imperial College London.
[38] García-Ordás, María Teresa, Enrique Alegre, Víctor González-Castro, and Rocío Alaiz-Rodríguez, *A computer vision approach to analyze and classify tool wear level in milling processes using shape descriptors and machine learning techniques.* The International Journal of Advanced Manufacturing Technology, 2017. **90**(5-8): p. 1947-1961.