July 22, 2021

Research on object detection and recognition using machine learning algorithms and deep learning: Approaches, Challenges, and Future Directions

Yousef Elbayoumi

Abstract

This article provides a thorough examination of machine learning and deep learning applications in autonomous vehicles, image detection, and credit card fraud detection. Deep learning is one potential answer for object identification and scene perception difficulties, allowing algorithm- and data-driven autos to be developed. Furthermore, utilizing image processing to accomplish automated fruit identification or recognition is a critical part in precision agriculture for performing object detection in big crop plots. In addition, this article discusses the identification of financial fraud in unbalanced data. I will evaluate and contrast various approaches to the problem of credit card fraud detection. Finally, the paper will go over some details about industry 4.0, as well as a poll that was completed on object detection.

Keywords:

Object Detection, Machine Learning, Industry 4.0, Self-Driving Cars, Deep Learning, Autonomous Driving Initiatives, Pineapple Crown, Image Processing, Precision Agriculture, Internet of Multimedia Things, Event Processing, Fraud detection, Poisson Process, Imbalanced Data.

1. Introduction

All previous estimates (earlier predictions) are being scaled back due to the explosive expansion in the number of physical devices connected to the Internet. We can expect billions of connected objects in the coming year as the next 10 billion Internet of Things (IoT) gadgets come online, and the prior "50 billion devices" figure by 2020 will only be well-cited [3]. The influence of this increase can also be seen on the worldwide network in the traffic of multimedia, with Netflix 14.97 percent, YouTube 11.35 percent, HTTP media stream 13.07 percent, and so on accounting for 57.69 percent of total Internet traffic [7]. Moreover, machines, products, and humans are increasingly connected via information technology, and the sector is undergoing a change toward more autonomous and intelligent manufacturing. The Fourth Industrial Revolution (Industry 4.0) is the name given to this process [5].

During this revolution, various applications of artificial intelligence (AI), machine learning (ML), and deep learning (DL) have acquired prominence and come to the fore as a result of recent developments in these techniques. Self-driving cars are one such application that is expected to have a profound and revolutionary impact on society and the way people commute [8]. Despite the fact that acceptance and domestication of technology can be difficult at first, these cars will be the first significant integration of personal robots into human civilization [2].

Another usage of object detection is detecting images, precision agriculture has lately gained traction as a result of the implementation of a computational analysis using image processing to aid the agricultural management team in monitoring, measuring, and responding to crop variability for better farm-level management. Farmers could profit greatly from image processing in precision farming or agricultural applications for various crops using automatic detection and fruit yield counts during harvesting season [13].

The final thing I included to this research is related to banks, most banks now offer secure internet services to their customers. A system for detecting and preventing fraudulent transactions is one of the components of such protection. The Poisson process intensity model and supervised machine learning methods are used in this study to solve the fraud detection problem. For that topic, various unsupervised techniques are employed, with the Restricted Boltzmann Machine (RBM) and Generative Adversarial Networks (GAN) being highlighted [1].

2. Background

The authors employ a physical Fischer Technik (FT) industrial simulation model to simulate an Industry 4.0 production environment1. Learning Factories [22] are models like this that are utilized for education and Industry 4.0 research [26]. This enables the development of research prototypes and the assessment of their suitability for application in practice. Authors believed that physical simulation models are closer to the real world than solely virtual simulation models (e.g., based on Digital Twins [18], notably in terms of developing runtime behavior and unanticipated ad-hoc interactions (e.g., by humans) not explored in virtual models [27].

2.1. Self-Driving Vehicles

The concept of self-driving automobiles has been around for about 80 years, with General Motors' (GM) Futurama [19] presenting it at the 1939 World's Fair in New York. The emergence of accurate and resilient sensors that continue to shrink in size and cost, along with AI, has been the cornerstone for autonomous driving systems [17]. Human-machine interface applications, network-enabled controls, multiple-sensor data fusion, 3D drive scene analysis, and software-defined signal processing are all included in these self-driving systems for transporting materials, payloads, goods, and people [14]. Self-driving machines based on AI must be able to navigate properly in any environment at any time [10]. Accurate localization, unobtrusive data gathering, fused data-set generation, and uninterrupted high-level communication with other cars and nearby smart infrastructure are all critical to autonomous navigation accuracy [20]. Self-driving technology is planned to be applied to tractor-trailers, cargo trucks, mining trucks, and buses in the long run [16].

Advantages of Self-Driving Vehicles:

Intelligent transportation systems (ITS) use advances in wireless networking, software-defined networking, and information and communication technology (ICT) to reduce collisions, reduce pollution, alleviate mobility issues, provide newer modes of public transportation, and share resources, materials, and space [4]. According to research, 1.3 million people die each year as a result of intoxicated, drugged, distracted, or tired driving. These lives could be saved if autonomous AI systems could eliminate some of the human follies [15]. The following benefits motivate current self-driving automobile research:

- Users may benefit from reduced stress, shorter commutes, shorter travel times, increased productivity, optimal fuel consumption, and lower carbon emissions. These vehicles can be programmed to drive cautiously, avoid blind areas, and adhere to speed regulations [6].
- Self-driving cars would help governments with traffic enforcement, increase roadway capacity, reduce road fatalities and the frequency of on-road driving-related accidents, and improve speed limit observance [11].
- Self-driving cars are expected to eliminate drunk driving, distracted driving, texting, and other forms of mobile phone use, as well as less braking and acceleration and highway jams [25]. Reduced accidents are predicted to benefit youngsters and the elderly, helping people to feel at ease with self-driving vehicles [29].

2.2. Image Detection

Pre-processing, segmentation, and classification are required for the processing of images acquired by an Unmanned Aerial Vehicle (UAV) at a given height in order to improve image representation or object recognition. Image processing based on automated disease identification, crop stress monitoring, yield prediction, and machine counting [21] have all been examined for their usefulness in crop yield maximization and productivity management. The rapid development of computer processing technology, which gives benefits in terms of quick deployment, low cost, and precise processing outcomes for huge in-field areas, is the cause for pertinence in numerous aspects of the agricultural sector.

Machine learning has a lot of potential for speedy and reliable results, which will be tested with relevant characteristics using analysis of variance (ANOVA) to improve the efficiency of the algorithm. To improve the detection process, high-quality photos with well-lit backgrounds and a proper segmentation approach are necessary [13].

2.3. Credit Card Fraud Detection

Working with a highly unbalanced sample makes detecting fraudulent transactions much more difficult, as few examples of the minority learning class are incorrectly detected by classifiers. To overcome this challenge, several techniques will be developed [9]. I will be using mathematical concepts and variation within the Poisson process to build the model, for more details on the mathematical part and how the model was built see the article [12].

Several machine learning algorithms will be used in order to achieve this problem, such as LightGBM, XGBoost, and CatBoost, also after data preprocessing, we will see many interesting numbers and results, and finally we will go over the computational process [12].

3. Methodologies

3.1. Object Detection Approaches in Self-Driving Vehicles: Pros and Cons

Object detection is a technique for determining which class instances an object belongs to [28]. Self-driving cars must classify different items present in an image, as well as the specific locations of these objects, in order to obtain a comprehensive 3D perspective of the surroundings [37]. The following three categories can be used to categorize object detection for semantic scene understanding:

1. **Region proposal/region selection:** in the pre-DL era, the most common method for region selection was to scan the entire image using a multi-scale sliding window. The sliding window technique, on the other hand, is computationally costly for self-driving automobiles and fails to meet the real-time requirement of exhaustively finding all positions of an item [24].
2. **Feature extraction:** for feature extraction, techniques such as the Haar-transform, Haar-like features, and histograms of oriented gradients (HoG) were commonly

utilized. These approaches, on the other hand, do not give robustness to changing environmental variables in self-driving scenarios [44].

3. **Categorization:** after the items have been perceived and located, classification is carried out using machine learning techniques like MLP and SVM. The Deformable Parts Model (DPM) has been proposed in self-driving situations [23] and has garnered universal support for object classification.

The histogram of oriented gradients (HoG) is a feature detection method that has been used to detect pedestrians. The histogram of oriented gradients produces gradients at various scales for the entire image and uses a linear classifier for each scale at each pixel. To avoid crashes, the HoG model for self-driving cars in congested areas with several vehicles requires near-perfect precision. For real-life circumstances including pedestrians and other cars, the model proved to be too slow [31]. Although several object detection architectures including as R–CNN, fast R–CNN, faster R–CNN, YOLO, and SSD have achieved impressive accuracy and low error rates (less than 5%) on ImageNet and Pascal VOC datasets, the speed of these designs in real-time self-driving applications remains a challenge [35].

3.2. Image Detection Approaches

Following the segmentation process, features are extracted and selected in this section. The appearance of the pineapple crown photos is represented by color features, and shape and texture features are extracted using geometric characteristics. The crown of the fruit is then distinguished from background noise using an analysis of these properties. Three major feature vectors were developed and recognized based on the sample photos of this framework: Color features, Shape features, and Texture features [13].

1- **Color features:** a color histogram on the pineapple crown was computed for the three matrices of R, G, and B to extract the shape aspects of the pineapple crown from the segmentation procedure. The mean R, mean B, and mean G are the three key features to be analyzed. Because of the background noise, color similarities amongst pineapple crowns were visible, indicating that minor variances existed [30].
2- **Shape features:** the size and shape of each pineapple crown, as well as background noise, are measured using shape characteristics. The area, centroid x, centroid y, major axis length, minor axis length, orientation, solidity, eccentricity, and perimeter are all components of the shape characteristic to be evaluated [30].
3- **Texture features:** mean LBP, standard deviation LBP, contrast, correlation, energy, mean GLCM, standard deviation GLCM, entropy, variance, kurtosis, smoothness, homogeneity, root mean square, and skewness are among the LBP and GLCM features investigated in texture extraction [36].

Object classification with machine learning

In general, classification in computer science and related fields enforces a computational model inspired by the central nervous system to handle non-linear issues corresponding to noisy or complicated data [32], including image analysis. Machine learning classifiers learn from morphological cues and count the number of fruit crowns in successive frames of images to estimate the number of pineapple crowns within the bounding box [34].

To classify the pineapple crown as a non-fruit crown, a machine learning classification technique is required to remove the genuine positive detection. The image algorithm-based pre-processing does not detect all of the bounding boxes. As a result, implementing a classification algorithm improves the system's ability to correctly identify the crown, making it beneficial during the fruit counting process [13].

Super Vector Machine: SVM is a straightforward data classification method that produces a model that predicts the goal value of data in the testing set. The model is built using training and testing data, which is made up of a series of data examples with one target value and a variety of features [33].

Random Forest: the RF is one of the most effective classification and regression methods for classifying large datasets. An ensemble of decision trees is generated by this algorithm. Ensemble approaches are based on the idea of grouping weak learners together to generate a strong learner [42].

Decision Trees: DT are a widely used non-parametric supervised machine learning technique for data classification. The fundamental purpose of DT is to develop a model that predicts the class label of a test sample by learning some rules from the testing dataset [39].

K-Nearest Neighbors: KNN classifier is a well-known methodology in the machine learning domain. The KNN classifier uses the training examples, a distance parameter, and the number of nearest neighbors to assign class labels to the test data (k) [39].

3.3. Credit Card Fraud Detection Approaches

Ensembles in machine learning are a collection of algorithms that have been trained to solve the same issue. As a result, ensembles outperform each algorithm in the ensemble separately in terms of forecasting efficiency. Using an anti-gradient, several gradient boosting models are created consecutively. In order to eliminate mistakes, the models in the learning process suggest the direction of future corrections in the predictions of the current ensemble model [12].

LightGBM: this is a more advanced version of the gradient boosting algorithm. With a tree-based training technique, LGBM is employed to improve the gradient. The key distinction between this algorithm and others is that the tree grows in depth, rather than by leaves. You should also pay attention to this algorithm's name. ”Light” denotes a rapid rate of execution. LightGBM is capable of handling enormous volumes of data while requiring the least amount of memory. Another advantage of this algorithm is its concentration on forecast accuracy [38].

XGBoost: it’s a decision tree-based machine learning method. It is a system optimization and algorithm improvement that improves the gradient boosting framework. XGBoost can be used to tackle regression, classification, ordering, and custom prediction problems, among other things. XGBoost is a parallel tree boosting algorithm that solves a variety of data science tasks quickly and accurately [40].

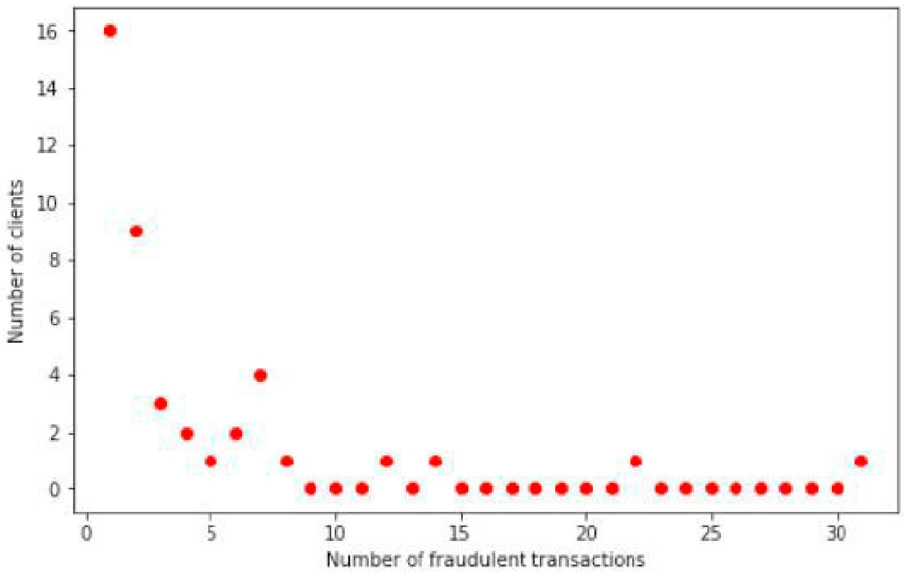
CatBoost: Yandex produced an open software library that uses one of the first gradient augmentation approaches to implement a unique patented algorithm for generating machine learning models. Numeric characteristics can be used in almost any current gradient-based approach. If the data set comprises both numerical and categorical signs, we must transform the categorical signs to numerical, which may result in a reduction in model accuracy. CatBoost is a gradient enhancement library having the advantage of being able to operate with two different sorts of characteristics [41].

Data preprocessing: between November 15, 2018, and February 13, 2019, 95 662 transactions were recorded. 3 633 clients have transaction data. This set's tuples all have a label indicating whether they belong to class 0 or 1. A strong data imbalance is an essential element of the data set: the percentage of fraudulent transactions is slightly less than 0.2 percent. To use the Poisson process basics to estimate the probability of an object belonging to a specific class, three attributes are needed: the client ID, the time the transaction arrived, and the label [12].

Based on the findings of data processing, the following statistics are provided [12]:

- 94 850 transactions.
- 2821 clients.
- 183 fraudulent transactions.
- 0.19% is percentage of fraudulent transactions.
- 42 fraudulent clients.
- 31 is maximum amount of fraud per client, and 0 is the minimum.

Computational Process Result



4. Results and Conclusion

4.1. Self-Driving Vehicles

The authors discussed researchers' continued efforts to test self-driving cars, emphasizing the need of deep learning in real-time object identification. DL could process acquired data in real-time and communicate it to neighboring clouds and other vehicles in the meaningful neighborhood, thanks to GPU and cloud-based fast computing. Transfer learning is also employed to improve accuracy of object detection, according to the study, in order to improve performance measures like as accuracy, precision, recall, and F1 scores [35].

DL is a major catalyst for realizing object detection and scene understanding in self-driving cars, according to the research, but there is still a lot of room for improvement. When and under what situations CNNs stop performing adequately and become a threat to human life in self-driving scenarios has yet to be determined [43].

For object categorization, multimodal sensor fusion and point cloud analysis were used to improve the limited exposure of the self-driving LiDAR cameras. Self-driving cars are no longer an issue of if, but rather of when and how, according to the survey's conclusions. The rate at which these autonomous robots become integrated into human civilization is determined by their capacity to drive safely. This necessitates the use of trustworthy object recognition techniques, mathematical models, and simulations to simulate reality and determine the ideal parameters and configurations that can react to changes in the environment [35].

Nonetheless, with large data, deep learning, and CNNs, we have technologies at our disposal to tackle perception problems in self-driving cars with high degrees of arbitrary accuracy. Researchers have been able to break down big problems into simpler ones, and formerly unsolvable challenges into doable but slightly more expensive ones, such as acquiring and annotating data to generate ground truth, thanks to these technologies [43].

4.2. Image Detection

Due to occlusions from background noise, as well as color similarities between leaves and crowns, image analysis for pineapple crown detection in an open in-field setting necessitates methodical processing phases. Comparing classifiers revealed that FN and FP errors are still significant when classifying pineapple and noise, which could be due to the similarities in features of pineapple and noise, particularly the color and form of pineapple leaves in these cases. Because the processing image method can reduce illumination and distinguish occlusions, which are required for image analysis at this stage, the possibility for desired ROI identification is great [13].

Fruit counting, which is essential to estimate yield in each plot of pineapple plantation, requires accurate pineapple crown detection. The use of ANOVA to pick features from color, texture, and form has shown to improve classification accuracy and produce superior results. It's clear that not all of the extracted traits were important enough to proceed with the categorization. There are 22 features picked from a total of 26 by matching the important differences of the pineapple crown, and background noise could be decreased. Shape, color, and texture characteristics were reduced to four major features, eccentricity, homogeneity, root mean square, and meanG, which showed that the characteristics are not statistically different for both pineapple and noise [13].

Using machine learning methodologies, a system using a DJI Phantom 3 Advanced quadcopter and MATLAB software was employed to perform real-time automatic crop recognition, detection, and counting of pineapples. In real time, the density count of pineapples in the selected area is calculated. This research provides an image-processing and machine-learning method for precisely detecting and counting pineapple crowns. There are various steps in the proposed method. To begin, improve the data image's quality and use morphological techniques to detect the pineapple crown. Second, use color, shape, and texture data as input to a machine learning classifier to differentiate between the pineapple crown and background noise like leaves, grass, and ground [51].

Finally, the pineapple fruits must be numbered based on their recognized crown by an automatic counting algorithm to display pineapple yield, with the feasibility of the system proved through testing of unseen photographs. ANN-GDX machine learning algorithm resulted in 94.4 percent accuracy as the best classification when compared to other classifier algorithms, as validated by testing of unseen images and capacity to overcome issue of varying illumination and occlusion owing to background noise [54].

4.3. Credit Card Fraud Detection

Two approaches were used to solve the problem of fraud detection: the Poisson process and machine learning. In the first scenario, I looked at a variety of intensity functions that can be used to forecast fraudulent events. As machine learning techniques, the gradient boosters LightGBM, XGBoost, and CatBoost were utilized. There were other solutions to the problems of data imbalance, "False positive answers," and the presence of "clean" clients [12].

New data was processed and generated during the project. It is sufficient to know the deterministic intensity function, the arrival time of the fraud transaction, and the label to estimate the Poisson process intensity. All of the features in the new dataset were employed in gradient boosting models [12].

Poisson process models will be built using the sliding window method in the future; more complex intensity functions for non-homogeneous processes, such as the Fourier series, will also be investigated. Ensembles from the above ensembles will be included to the machine learning research. It is expected that by combining the results of developing such models with multiple ways for identifying fraudulent transactions, the disadvantages of one algorithm will be balanced by the benefits of another [46]. The strategy of applying the Poisson process to financial datasets, when combined with machine learning, can lead to the best appropriate solution for fraud detection.

5. Challenges and Future Directions

5.1. Challenges

Standardization of the idea IoMT: We discovered that existing IoMT-based applications are domain-specific, indicating that a standardized IoMT architecture is needed. [48] and [52] pioneered the standardization of architecture for the Multimedia Internet of Things, addressing concerns such as multi-modal Big data computation, as well as the model's scalability and maintainability for effective multimedia information exchange [52]. However, it is a new problem that will require more attention from various IoT and multimedia communities to agree on and address all of the requirements of IoMT-based smart city systems.

Multimedia-based IoT Approaches: We found that IoT infrastructures are well-established, crowded, and well-equipped with adequate communication and processing protocols, as well as optimization strategies. Inclusion of multimedia in IoT, on the other hand, is the greatest problem we have, owing to the lack of attention paid to this gap in both disciplines in the past [45]. Multimedia applications are solely concerned with feature extraction, processing, and analysis of multimedia data, whereas the Internet of Things is concerned with scalar (structured) data. We anticipate that low-cost IoT models will work for multimedia and that multimedia devices will use fewer resources. A critical but necessary problem to study is bringing both strong regions together and developing them from the ground up.

Communication Is Expensive: Some of the significant issues of IoMT-based data include high bandwidth requirements, excessive energy consumption, and relatively high heterogeneous multimedia data. A few recent models [47] also seek to meet the large bandwidth difficulty while reducing end-to-end delay, emphasizing their significance for multimedia traffic.

Due to the rise in the volume of multimedia big data (such as films and photos from smartphones), energy-efficient processing has become a top priority challenge in multimedia-based IoT [49]. Because of the growing number of multimedia applications (such as smart homes, transportation, security systems, and manufacturing), heterogeneity [50] is being used in IoT and is also viewed as the basic issue of future IoT.

5.2. Future Directions

Event-based middle wares are well-known IoT solutions that isolate the intricacies of the system/hardware from the application developer [53]. Multimedia aware Middle wares: Existing event-based middle wares have a rich literature for structured event processing and have succeeded to cause a revolution in distributed system communication architectures [63]. As a result, we can confidently predict the success of multimedia events-based middle wares for multimedia-based IoT solutions.

Unified Communication Paradigm: For the sending and receiving of structured events, event-based middle wares have so far agreed on the same communication paradigms (such as publish/subscribe). Subscribers express their interest in the detection of any event using subscriptions and are notified when this event occurs [55] while communicating through the publish/ subscribe models. In recent study, expressive power for users has gotten a lot of attention in the field of approximation event processing, with impressive results in message forwarding and redundancy removal [56]. Furthermore, adaptive filtering-based publish/subscribe for distributed multimedia-based systems showed a significant reduction in latencies and improved resource efficiency [57]. Another microservice design validates the reduction in energy consumption for multimedia event processing utilizing publish/subscribe communication. These advancements, we feel, are required for uniform architecture and successful transmission of multimedia events [58].

Multimedia Query Languages: In addition to structured event query languages, many video query languages have been proposed for the processing of image-based events, such as VEQL, CVQL, SPARQL-MM, SVQL [59]. The majority of these languages have the ability to detect multimedia items, support detection properties, anticipate spatial/temporal relationships, and process streams efficiently. I believe that adding multimedia query languages into real-time applications for the representation and specification of multimedia events could be extremely beneficial to smart cities [60]. Furthermore, the high throughput of query language-based architectures could have a significant impact on Multimedia-based IoT systems.

Models based on Deep Neural Networks (DNN): Deep learning has made substantial advances in the field of image recognition, paving the way for smart city surveillance applications [61]. Incorporating deep convolutional networks-based approaches for multimedia event analysis could be a potential answer for IoMT methodology standardization. The work [64] implements a generic technique for IoMT data and illustrates the ability of deep neural networks to process multimedia event streams from numerous applications. DNNs' capacity to give high performance and continuous learning can help to meet the demands of multimedia in IoT [62]. DNN-based techniques may incorporate any type of classifier to facilitate many types of applications in smart cities, regardless of whether they provide high-performance capabilities in image identification.

6. Acknowledgement

This research is hosted by Data Glacier, and I would like to thank my university, Bahçeşehir University for their great instructors who helped me to reach the point where I can write articles.

I used many references, but there was 5 main references which helped me a lot and I basically built this article according to them, I would like to thank the authors for their great work and research, these are the articles: [12], [13], [35], [65], and [66].

Declaration of Competing Interest

I declare that I have no known competing financial interests or personal ties that could have influenced the research reported in this paper.

References

[1] Niu, X., Wang, L. and Yang X. (2019). A comparison study of credit card fraud detection: Supervised versus Unsupervised. [online] arXiv:1904.10604 [cs.LG]. Available at: https://arxiv.org/abs/1904.10604.

[2] Kellett J, Barreto R, Hengel A, Vogiatzis N. How might autonomous vehicles impact the city? the case of commuting to central adelaide. Urban Pol Res 2019; 37:442–57.

[3] Amy Nordrum, The internet of fewer things [news], IEEE Spectr. 53 (10) (2016) 12–13.

[4] Divakarla K, Emadi A, Razavi S, Habibi S, Yan F. A review of autonomous vehicle technology landscape. Int J Electr Hybrid Veh (IJEHV) 2019;11:320–45.

[5] Lasi, H., et al., 2014. Industry 4.0. BISE 6, 239–242.

[6] Casey A, Niblett A. Self-driving contracts. J Corp Law 2017;43:1–33.

[7] Lyn Cantor, Cam Cullen, The global internet phenomena reportoctober 2018, Technical Report. (2020).

[8] Pakusch C, Stevens G, Boden A, Bossauer P. Unintended effects of autonomous driving: a study on mobility preferences in the future. Sustainability 2018;10: 2404.

[9] Wedge, R., Kanter, J.M. and Veeramachaneni, K. (2017). Solving the “false positives” problem in fraud prediction. [online] arXiv:1710.07709 [cs.AI]. Available at: https://arxiv.org/abs/1710.07709.

[10] Brell T, P R, Martina Z. Suspicious minds? - users’ perceptions of autonomous and connected driving. Theor Issues Ergon Sci 2019;20:301.

[11] Zhang B, Wilschut E, Willemsen D, Martens M. Transitions to manual control from highly automated driving in non-critical truck platooning scenarios. Transport Res Part F: Psychology and Behaviour 2019;64:84–97.

[12] Anastasiia I., Adel V., Comparison of Poisson process and machine learning algorithms approach for credit card fraud detection, Procedia Computer Science 186 (2021) 721–726.

[13] Wan Nurazwin S. R., Muhammad A. H., Megat S. A. M.A., Nur D. K.A., Automated image identification, detection and fruit counting of top-view pineapple crown using machine learning, Alexandria Engineering Journal (2021).

[14] Cho E, Jung Y. Consumers understanding of autonomous driving. Inf Technol People 2018;31:1035.

[15] de Winter J, Happee R, Martens M, Stanton N. Effects of adaptive cruise control and highly automated driving on workload and situation awareness: a review of the empirical evidence. Transport Res Part F: Psychology and Behaviour 2014;27: 196–217.

[16] Johannesson L, Murgovski N, Jonasson E, Hellgren J, Egardt B. Predictive energy management of hybrid long-haul trucks. Contr Eng Pract 2015;41:83–97.

[17] Santos V, A D S, Miguel O. Special issue on autonomous driving and driver assistance systems. Robot Autonom Syst 2017;91:208–9.

[18] Boschert, S., Rosen, R., 2016. Digital Twin—The Simulation Aspect, in: Mechatron. Futur.. Springer, pp. 59–74.

[19] Wasson H. The other small screen: moving images at New York’s world fair, 1939. Can J Film Stud 2012;21:81–103.

[20] Laes E, Gorissen L, Nevens F. A comparison of energy transition governance in Germany, The Netherlands and the United Kingdom. Sustainability 2014;6: 1129–52.

[21] S. Gutie´rrez, J. Tardaguila, J. Ferna´ ndez-Novales, M.P. Diago, On-the-go hyperspectral imaging for the in-field estimation of grape berry soluble solids and anthocyanin concentration, Aust. J. Grape Wine Res. 25 (1) (2019) 127–133, https://doi.org/ 10.1111/ajgw.12376.
[22] Abele, E., et al., 2017. Learning factories for future oriented research and education in manufacturing. CIRP Ann. 66, 803–826.
[23] Witoonchart P, Chongstitvatana P. Application of structured support vector machine backpropagation to a convolutional neural network for human pose estimation. Neural Network 2017;92:39–46.
[24] Li H. Zoom out-and-in network with map attention decision for region proposal and object detection. Int J Comput Vis 2018:1–14.
[25] Louw T, Markkula G, Boer E, Madigan R, Carsten O, Merat N. Coming back into the loop: drivers perceptual-motor performance in critical events after automated driving. Accid Anal Prev 2017;108:9–18.
[26] Klein, P., Malburg, L., Bergmann, R., 2019. FTOnto: A Domain Ontology for a Fischertechnik Simulation Production Factory by Reusing Existing Ontologies, in: Proc. of the Conf. LWDA, CEUR-WS.org. pp. 253–264.
[27] Broy, M., Cengarle, M.V., Geisberger, E., 2012. Cyber-Physical Systems: Imminent Challenges, in: Large-Scale Complex IT Syst. Dev., Operat. and Manag. - 17th Monterey Workshop, Springer. pp. 1–28.
[28] X. Chen, Monocular 3d object detection for autonomous driving, DOI: 10.1109/ CVPR.2016.236.
[29] Gish J, Grenier A, Vrkljan B, Van Miltenburg B. Older people driving a high-tech automobile: emergent driving routines and new relationships with driving. Can J Commun 2017;42:235.
[30] B.S. Vidya, E. Chandra, Entropy based Local Binary Pattern (ELBP) feature extraction technique of multimodal biometrics as defence mechanism for cloud storage, Alexandria Eng. J. 58 (1) (2019) 103–114, https://doi.org/10.1016/j.aej.2018.12.008.
[31] Watanabe T, Ito S, Yokoi K. Image feature descriptor using co occurrence histograms of oriented gradients for human detection. J Inst Image Inf Televis Eng 2017;71:J28–34.
[32] H.A. Babikir, Elaziz M. Abd, A.H. Elsheikh, E.A. Showaib, M. Elhadary, D. Wu, Y. Liu, Noise prediction of axial piston pump based on different valve materials using a modified artificial neural network model, Alexandria Eng. J. 58 (3) (2019) 1077– 1087, https://doi.org/10.1016/j.aej.2019.09.010.
[33] R. Alzu'bi, A. Anushya, E. Hamed, E.A. Al Sha'ar, B.A. Vincy, Dates fruits classification using SVM, in: AIP Conf. Proc., AIP Publishing LLC, 1952(1), 2018, pp. 20078.
[34] P. Roy, A. Kislay, P.A. Plonski, J. Luby, V. Isler, Vision-based preharvest yield mapping for apple orchards, Comput. Electron. Agric. 164 (2019), https://doi.org/10.1016/j.compag.2019.104897 104897.
[35] Abhishek G., Alagan A., Ling G., Ahmed S. K., Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues, Array 10 (2021) 100057.
[36] R.B. Vallabhaneni, V. Rajesh, Brain tumour detection using mean shift clustering and GLCM features with edge adaptive total variation denoising technique, Alexandria Eng. J. 57 (4) (2018) 2387–2392, https://doi.org/10.1016/j.aej.2017.09.011.
[37] Hane C. 3d visual perception for self-driving cars using a multi camera system: calibration, mapping, localization, and obstacle detection. Image Vis Comput 2017;68:14–27.
[38] LightGBM (2020). Documentation LightGBM. [online] LightGBM. Available at: https://lightgbm.readthedocs.io/en/latest/.

[39] A. Al-Zebari, A. Sengur, Performance Comparison of Machine Learning Techniques on Diabetes Disease Detection, in: 2019 1st Int. Informatics Softw. Eng. Conf., IEEE, 2019, pp. 1–4.
[40] XGBoost (2020). Documentation XGBoost. [online] XGBoost. Available at: https://xgboost.readthedocs.io/en/latest/.
[41] CatBoost (2020). Documentation CatBoost. [online] CatBoost. Available at: https://catboost.ai/.
[42] E. Elhariri, N. El-Bendary, A.E. Hassanien, A. Badr, A.M. Hussein, V. Sna´ sˇel, Random forests based classification for crops ripeness stages, in: Proc. Fifth Int. Conf. Innov. Bio-Inspired Comput. Appl. IBICA 2014, Springer, 2014, pp. 205– 215.
[43] Dairi A, Harrou F, Senouci M, Sun Y. Technology - cybernetics; new cybernetics findings from sejong university described (efficient deep cnn-based fire detection and localization in video surveillance applications). J Technol 2019:1207.
[44] Wang X. Regionlets for generic object detection. IEEE Trans Pattern Anal Mach Intell 2015;37:2071–84.
[45] Sheeraz A. Alvi, Bilal Afzal, Ghalib A. Shah, Luigi Atzori, Waqar Mahmood, Internet of multimedia things: Vision and challenges, Ad Hoc Netw. 33 (2015) 87–111.
[46] Zojaji, Z., Atani, R.E. and Monadjemi, A.H. (2016). A survey of credit card fraud detection techniques: data and technique oriented perspective. [online] arXiv:1611.06439 [cs.CR]. Available at: https://arxiv.org/abs/1611.06439.
[47] Ayoub Bahnasse, Fatima Ezzahraa Louhab, Hafsa Ait Oulahyane, Mohamed Talea, Assia Bakali, Novel sdn architecture for smart mpls traffic engineering-diffserv awaremanagement, Future Generation Computer Systems, 87, 2018, pp. 115–126.
[48] Kah Phooi Seng, Li-Minn Ang, A big data layered architecture and functional units for the multimedia internet of things (miot), IEEE Trans. Multi-Scale Comput. Syst. 4 (4) (2018) 500–512.
[49] Huy Trinh, Prasad Calyam, Dmitrii Chemodanov, Shizeng Yao, Qing Lei, Fan Gao, Kannappan Palaniappan, Energy-aware mobile edge computing and routing for low-latency visual data processing, IEEE Trans. Multimedia 20 (10 (2018) 2562–2577.
[50] Tie Qiu, Ning Chen, Keqiu Li, Mohammed Atiquzzaman, Wenbing Zhao, How can heterogeneous internet of things build our future: a survey, IEEE Commun. Surv. Tutorials 20 (3) (2018) 2011–2027.
[51] E. Miriti, Classification of selected apple fruit varieties using Naive Bayes, 2016.
[52] Sufyan Almajali, I. Dhiah el Diehn, Haythem Bany Salameh, Moussa Ayyash, Hany Elgala, A distributed multi-layer mec-cloud architecture for processing large scale iot-based multimedia applications, Multimed. Tools Appl. (2018) 1 22.
[53] Edward Curry, Message-orientedmiddleware,Middleware Commun. (2004) 1–28.
[54] R. Alzu'bi, A. Anushya, E. Hamed, E.A. Al Sha'ar, B.A. Vincy, Dates fruits classification using SVM, in: AIP Conf. Proc., AIP Publishing LLC, 1952(1), 2018, pp. 20078.
[55] Patrick Th. Eugster, Pascal A. Felber, Rachid Guerraoui, Anne Marie Kermarrec, The many faces of publish/subscribe, ACM Comput. Surv. (CSUR) 35 (2 (2003) 114–131.
[56] Niki Pavlopoulou, Edward Curry, Towards a window-based diverse entity summarisation engine in publish/subscribe systems, Proceedings of EYRE 19: 2nd International Workshop on Entity Retrieval, ACM, 2019.
[57] Tarek Zaarour, Niki Pavlopoulou, Souleiman Hasan, Umair Ul Hassan, Edward Curry, Automatic anomaly detection over sliding windows: Grand challenge, Proceedings of the 11th ACM International Conference on Distributed and Eventbased Systems 2017, pp. 310–314.

[58] Felipe Arruda Pontes, Edward Curry, Cloud-edge microservice architecture for dnn-based distributed multimedia event processing, EdgeWays 2020, Accepted for Springer's Communications in Computer and Information Science (CCIS), 2020.
[59] Tony C.T. Kuo, Arbee L.P. Chen, Content-based query processing for video databases, IEEE Trans. Multimedia 2 (1) (2000) 1–13.
[60] Piyush Yadav, Dipto Sarkar, Dhaval Salwala, Edward Curry, Traffic prediction framework for openstreetmap using deep learning based complex event processing and open traffic cameras, arXiv Preprint (2020) arXiv:2008.00928.
[61] MinWu,Marta Kwiatkowska, Robustness guarantees for deep neural networks on videos, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2020, pp. 311–320.
[62] Hao Li, Hong Zhang, Xiaojuan Qi, Ruigang Yang, Huang Gao, Improved techniques for training adaptive deep networks, Proceedings of the IEEE International Conference on Computer Vision 2019, pp. 1891–1900.
[63] René Meier, Vinny Cahill, Taxonomy of distributed event-based programming systems, Comput. J. 48 (5) (2005) 602–626.
[64] Asra Aslam, Edward Curry, Towards a generalized approach for deep neural network based event processing for the internet of multimedia things, IEEE Access 6 (2018) 25573–25587.
[65] Lukas M., Manfred-Peter R., Ronny S., Patrick K., Ralph B., Object Detection for Smart Factory Processes by Machine Learning, Procedia Computer Science 184 (2021) 581–588.
[66] Asra A., Edward C., A Survey on Object Detection for the Internet of Multimedia Things (IoMT) using Deep Learning and Event-based Middleware: Approaches, Challenges, and Future Directions, Image and Vision Computing 106 (2021) 104095.