International Journal of Applied Engineering Research

Identification and Detection of Objects Using Deep Learning Techniques

M.L.S.N.S Lakshmi[1], Dr.Vanajaroselin 2, S.Anitha[3], Sreevardhan Cheerla[4],

[1]*Associate professor, ECE Department QIS college of Engineering and Technology,Ongole,India*
E-mail Id:mlakshmi.0290@gmail.com
[3]*Assistant professor, ECE Department QIS college of Engineering and Technology,Ongole,India*
[4]*Associate professor Department of Electronics and Communication Engineering, KoneruLakshmaiah Education Foundation, Vaddeswaram, Guntur (Dist), Andhra Pradesh - 522302*
2 Professor, Department of ISE, The Oxford College of Engineering , Bangalore

detect cases of diverse objects in images, recordings, or video accounts is made possible by object discovery. It recognizes the various attributes of Pictures instead of item location strategies and produces a clever and powerful comprehension of pictures definitely like human vision works. As we move on, we'll start with a brief demonstration of deep learning and well-known article recognition frameworks including CNN, R-CNN, RNN (Recurrent Mind Organization), Faster RNN, and YOLO are examples of convolutional neural networks (You Only look once).From the proposed object recognition model engineering along with specific headways and adjustments. The traditional model perceives a little object shown in photographs. Our suggested model provides the proper result with accuracy. In this work we have performed the detection of objects using YOLO technique and accurate results have been achieved with comparison to R-CNN technique.

Keywords: Deep learning, CNN (Convolutional Neural Network),object detection,R-CNN, YOLO.

I. Introduction

Object detection has long been a fascinating study area and a main area of interest in computer vision, with applications in robotics, autonomous vehicles, video surveillance, and pedestrian identification. Deep learning technology has changed the conventional approaches to object detection and identification. The deep neural network is frequently used as the feature extraction module in object detection in image processing because it has a strong feature depiction capability. Using deep learning models doesn't call for any manually constructed features and may be used as both a classifier and a regression tool. Therefore, the future of technology for deep learning in the field of object detection is bright. The objective of object detection is to recognize and locate where items are actually placed in an image (object localization). The pipeline of traditional object recognition algorithms consists of the three steps of informative feature extraction and region selection, and detection.

1. Informative region selection: Because various items appear in different places and varying aspect ratios or sizes, the picture may be warped. A sliding window with many scales must be used to scan the full image. This exhaustive method has obvious shortcomings even if it can find every possible location for the items. Due to the high number of candidate windows, it generates numerous redundant windows and calls for a lot of processing. If only there was a chance of unsatisfactory portions in the offered image.

2. Feature Extraction: We must take out visual components that can provide a trustworthy and meaningful depiction if we are to distinguish between distinct products. This is owing to qualities in the human brain that can form representations associated with complicated cells.It is challenging directly to create a comprehensive attribute description that is able to precisely classify all types of objects due to the huge variation of looks, backgrounds, and lighting circumstances.

3. Detection: In order to improve the hierarchical and informative nature of the visual recognition presentations, a detector divides a target object into its own category.

The Supported Vector Machine (SVM) and the Models with Deformable Parts include typically appropriate options (DPM). Among these classifiers, the DPM is a versatile model that handles severe deformations by fusing object components with deformation cost. In DPM, low-level characteristics that have been thoughtfully created, and part decompositions motivated by kinematics are merged in a graphical model, and learning to discriminate between graphical models allow for development of models for a high-precision part-based variety of object classes.

The proposed model improves the accuracy of weak feature objects in complicated settings by leveraging their semantic relevance[1-5] focuses on the framework design and operational idea of the model, as well as its real-time functionality and detection precision [6–9] presents an easy-to-understand a simple, yet effective, formulation of the issue of object detection using bounding box masks for objects. It offers a various scales of inference method that employs two or three network applications to generate detections of objects in high resolution at an affordable price [10-14] presents an overview of deep learning with a focus on object recognition, detection, and segmentation, which are fundamental difficulties in computer vision with numerous applications to photos and videos. provides a method for completing jobs in a network of non overlapping multiple cameras [15-18]. There is vision employing a new object identification technique Objects are further divided based on mean shift (MS) segmentation and depth of knowledge obtained from templates for stereo fixed-number sliding windows. It's also feasible to use supervised learning to solve the problem by employing Decision trees or, more likely, SVM in deep learning, as demonstrated in[19,20] deals with the topic of computer vision, mostly for object detection tasks using deep learning. A short overview of the datasets and deep learning algorithms used in computer vision may be found here. presents a thorough examination of Frameworks for deep learning-based object recognition that deal with numerous sub-issues, like clutter and low resolution, to differing percentages of R-CNN modification [21-24] deals with the Easynet model, which allows for detection predictions with a single network. The Easynet model checks the full image during testing, so the forecasts are informed by the general environment [25]. Focuses on vehicle identification and recognition from a video feed. This approach produces higher outcomes in terms of precision, detection, and classification, with a 99.2% accuracy rate.

The image in this piece is originally divided into a number of grids. Each grid measures S × S. The grids created from an input image are displayed in the following graph. In order to determine an object's height, width, center, and class, YOLO employs a single bounding box regression to determine the height, breadth, centre, and class of an item. Box overlapping can be explained by the phenomenon of object detection known as intersection over union (IOU). YOLO makes use of IOU to produce an output box that precisely encloses the objects. The projected bounding boxes' confidence ratings and their responsible for a grid cell each. If the predicted and actual boundary boxes line up, the IOU is equal to 1. By using this technique, bounding boxes that differ from actual ones are eliminated.

II. Research Elaboration

Artificial intelligence's deep learning capability mimics how the human brain gathers information, creates patterns, and makes decisions. It is a subfield of artificial intelligence's machine learning that employs networks to learn unsupervised from data.

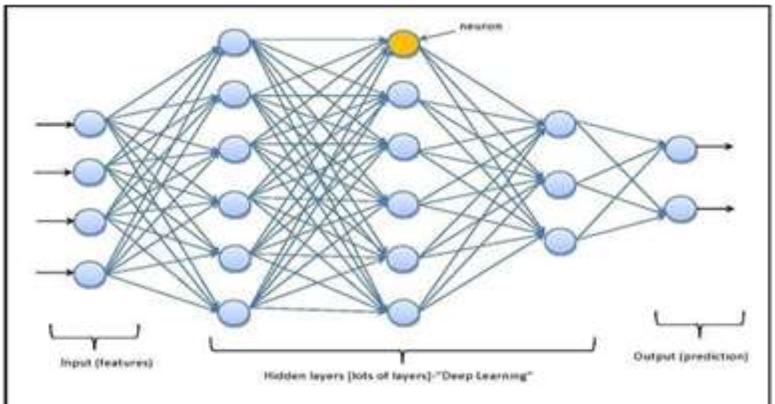

Figure.1ArchitectureofDeepLearning

A.CNN (Convolutional Neural Network): The three layers of a CNN are the input layer, the output layer, and numerous concealed layers. Convolution layers that combine with a dot product or multiplication make up a CNN's hidden layers. RELU layers are frequently additional convolutions like pooling layers, entirely linked layers, and normalising layers. Because their final convolution and activation function conceal their inputs and outputs, these layers are known as hidden layers. Typically, the product is a cross-correlation or sliding dot. These layers from CNN convolute the input before sending the output to the following layer.

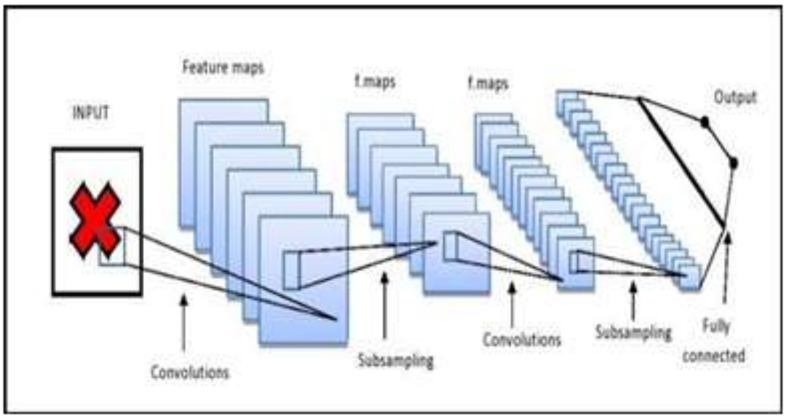

Figure.2 CNN

B. YOLO: In the YOLO (You Only Look Once) system, deep learning and convolutional neural networks (CNN) are utilized to recognize objects. There can only be one "seen" of each image. Because of this, YOLO is one of the fastest detection methods. It has a real-time detection rate of 30 frames per second. The image is split into a S*S grid for detecting purposes (left image). Each cell will anticipate N different bounding boxes and their probability levels. S*S*N boxes are calculated in this manner.

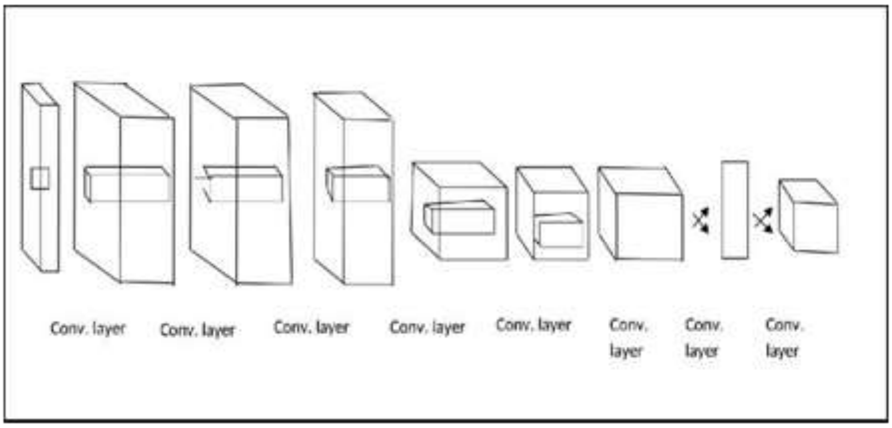

Figure.3YOLO

C.R-CNN: Region-based CNNs are what the R-CNN stands for. The R-CNN model chooses several suggested areas from a given image and then labels the bounding boxes and categories of those regions. It performs forward computation using a CNN to extract features from each proposed area. Then we estimate the categories and bounding boxes for each suggested region based on its attributes.

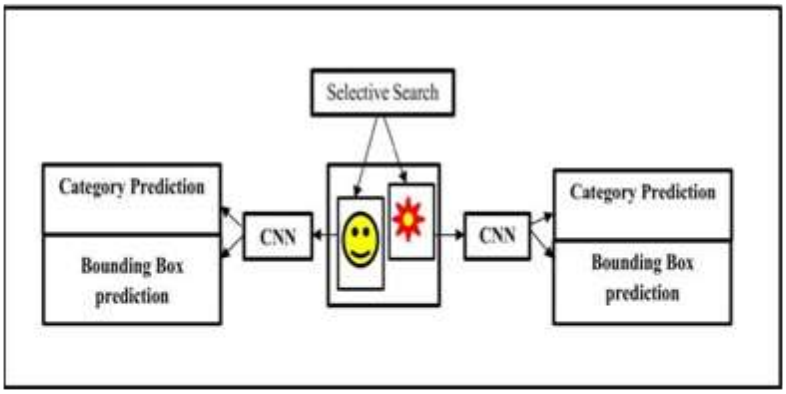

Figure.4R-CNN

III. Methodology

The following three methods are necessary for the YOLO algorithm to function: Bounding box regression, residual blocks, and intersection over union are the first three concepts (IOU).

Residual blocks

Several grids are first placed over the image. The sizes of each grid are S × S. The following graphic shows the grids that were produced using an input image.

Figure.5 Remaining blocks

There are several cells in the grid with the same size in the preceding picture. Objects that enter a grid cell will be able to be recognized by it. For instance, if the center of a thing is located in a particular grid cell, that cell will be responsible for identifying the object.

Bounding box regression

An outline that draws attention to an object in a picture is called a bounding box. The following characteristics are present in each bounding box in the image: Height, Class, Width, and Boxing are all in the middle (bx, by)

An illustration of a bounding box can be seen in the image below. A yellow outline serves as a representation of the bounding box.

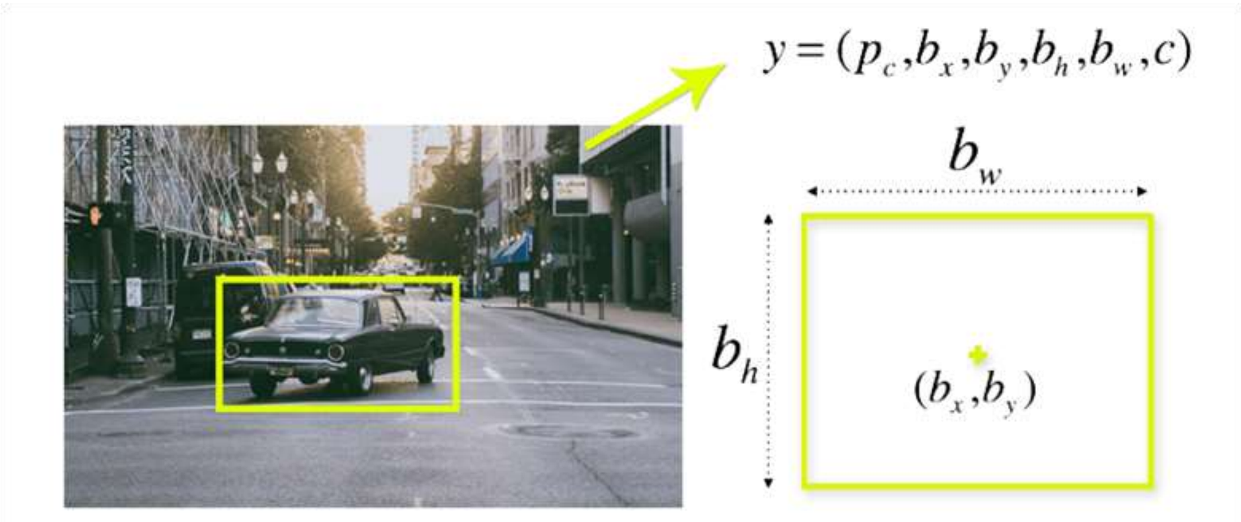

Figure.6 Bounding Box

As a way to ascertain an object's height, width, centre, and class, YOLO uses a single regression using bounding boxes demonstrates the possibility that an object will be located within the bounding box of the previous image.

Intersection over union (IOU)

Box overlapping can be explained by the object detection phenomenon known as intersection over union (IOU). To create a proper output container for the items. YOLO uses IOU.

a grid cell each is determined based on the anticipated bounding boxes' confidence ratings. If the planned and real bounding boxes line up, the IOU equals one. Boxes that are bound different from the real box are removed using this technique.

The simple operation of IOU is illustrated in the following figure.

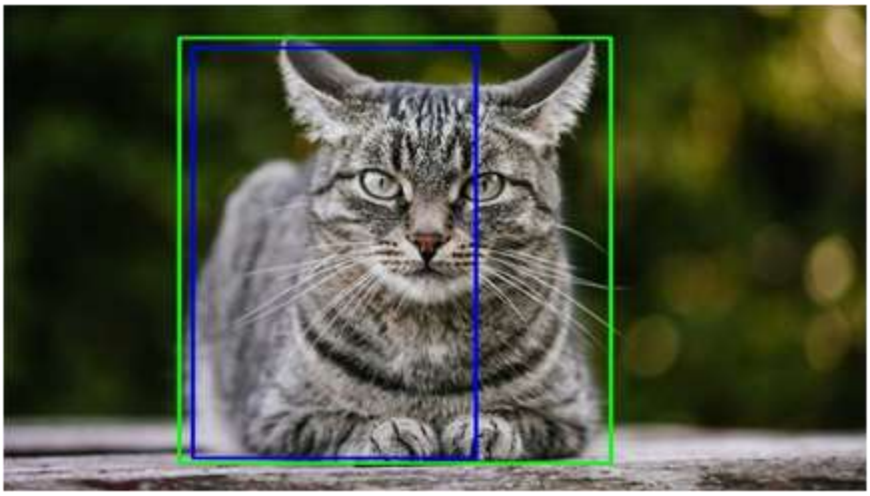

Figure.7 Intersection over union(IOU)

The above picture shows green and blue boundary boxes, respectively. The actual package is green, while the anticipated box is blue. Because of YOLO, both the bounding boxes are balanced.

Merging of the three techniques

The final detection findings, which may be seen in the accompanying image, are produced by combining the three techniques.

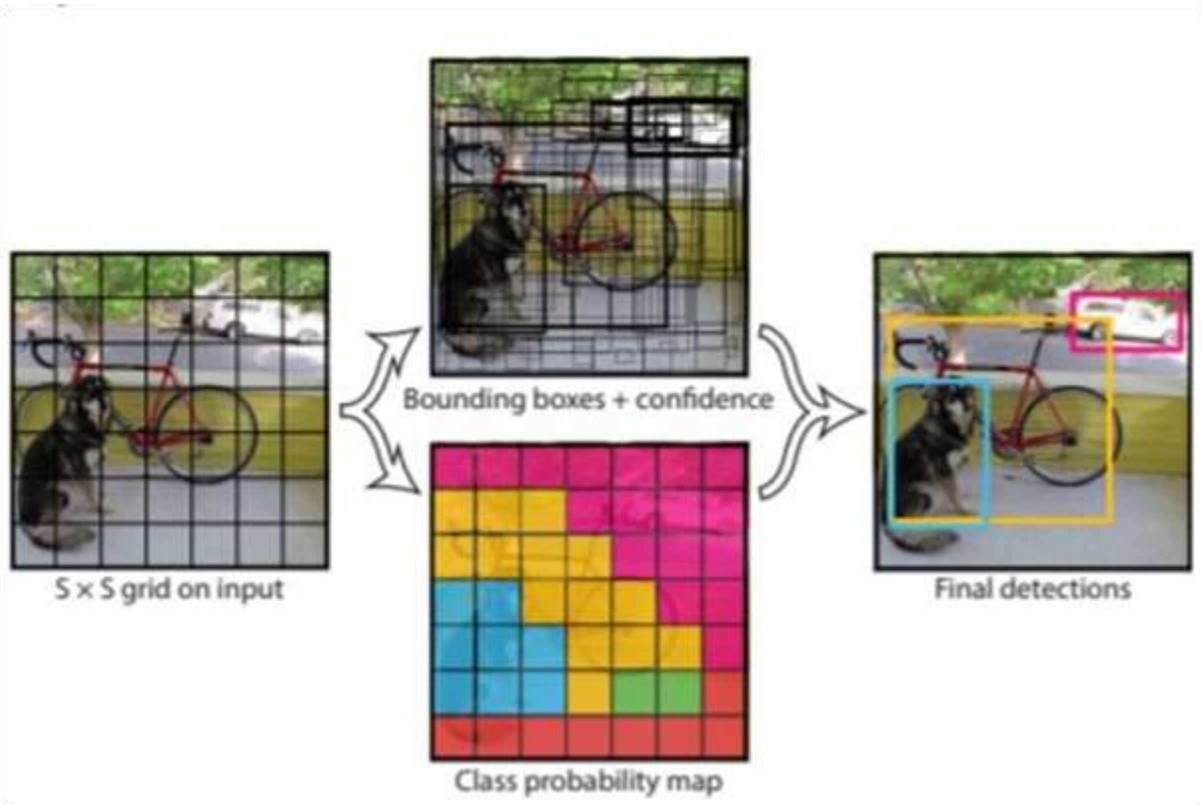

Figure.8 Combination of three techniques

Grid cells are initially created from the image. Every grid cell, B boundary boxes and confidence results are anticipated. The cells project the class probability to identify each object's class. A bicycle, a dog, and a car are three instances of objects that can be of at least three different sorts. One convolutional neural network makes all predictions simultaneously.

The projected bounding boxes and the actual boxes of the objects match because intersection triumphs over union. This occurrence causes bounding boxes that are extra or don't fit the items' attributes to be eliminated (like height and width). The ultimate detection will include special bounding boxes made especially to fit the items. For instance, the car is enclosed in the pink bounding box, whereas the bicycle is enclosed in the box with yellow borders. The dog has been highlighted using the blue bounding box.

Output

The output image displays the model's output. The output is executed using the fully linked layer as a foundation.

IV. Result

The precision and recall of an object identification model are used to assess how well it performs for each of the best-matched bounding boxes for the image's known objects.

Fig.9 Results of object detection when there are vehicles

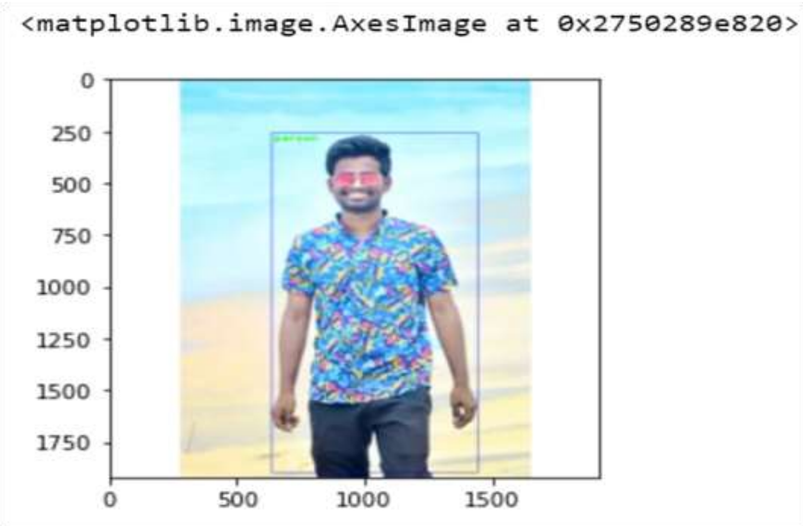

Fig.10 Results of object detection when a person

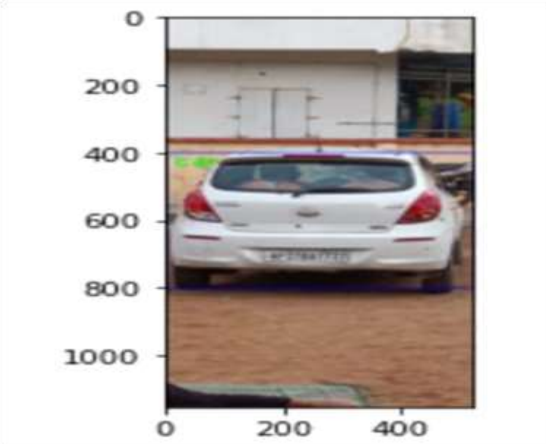

Fig.11 Object detection result in case of vehicle

V. Conclusion

As Object identification and One of the hardest things is recognition, intricate, and crucial problems in the field of computer vision today. As far as we are aware, the main objective of this project was to record real-time items in pictures, videos, or webcams.

• Pre-processing techniques are suggested here. Specifically, edge detection approaches that boost the image's contrast and raise the accuracy of our model.

• In the future, everyone can innovate and make improvements without concern for complexity.

• The project can be run on a GPU-equipped system for faster results and greater precision in the future.

• For instance, MS COCO does small object detection in several face detection tasks and applications. for better small object localization across partial barriers. Consequently, we will make certain changes to the network design to improve it.

• Implementing that enables us to minimise the burden on the data network.

• For more accurate and effective recognition of small things.

The final conclusion is that using picture enhancing and edge detection as pre-processing methods Utilizing augmentation and contrast will improve performance and accuracy.

References

[1] Ce Li, Yachao Zhang and YanyunQu, "Object DetectionBasedonDeepLearningofSmallSamples,"InternationalConference, pp.1-6,March2018.

[2] Cong Tang, YunsongFeng, Xing Yang, Chao Zheng andYuanpuZhou,"TheObjectDetectionBasedonDeepLearning,"InternationalConference, pp.1-6,2017.

[3] ChristianSzegedy,AlexanderToshevandDumitruErhan,"DeepNeuralNetworksforObjectDetection,"IEEE,pp.1-9,2007.

[4] Xiaogang Wang, "Deep Learning in Object Recognition,Detection,andSegmentation,"IEEE,pp.1-40,Apr.2014.

[5] Shuai Zhang, Chong Wang and Shing-Chow Chan, "NewObjectDetection,Tracking,andRecognitionApproachesforVideoSurveillanceOverCamera Network," IEEE SENSORS JOURNAL, vol. 15, no.69, pp.1-13,May2015.

[6] Malay Shah and Prof. RupalKapdi, "Object DetectionUsingDeepNeuralNetworks,"InternationalConference,IEEE,pp.1-4,2017.

[7] XiaoMa,KeZhouandJiangfengZheng,"Photo-RealisticFaceAgeProgression/RegressionUsingaSingleGenerativeAdversarialNetwork,"Neurocomputing,ElsevierB.V.,pp.1-16,July2019.

[8] Zhong-Qiu Zhao, PengZheng, Shou-taoXu and XindongWu,"ObjectDetectionwithDeepLearning:AReview,"IEEE,pp.1-21,2019.

[9] SandeepKumar,AmanBalyanandManviChawla,"Object Detection and Recognition in Images," IJEDR,pp.1-6,2017.
[10] Adami Fatima Zohra, SalmiKamilia, Abbas Faycal andSaadiSouad, "Detection and Classification Of VehiclesUsingDeepLearning,"InternationalJournalofComputer Science trends and technology(IJCST), vol.6,pp.1-7,2018.
[11] T. Guo, J. Dong, H. Li, and Y. Gao, "Simple convolutional neural network on image classification," 2017 IEEE 2nd Int. Conf. Big Data Anal. ICBDA 2017, pp. 721– 724, 2017, doi: 10.1109/ICBDA.2017.8078730.
[12] J. Du, "Understanding of Object Detection Based on CNN Family and YOLO," J. Phys. Conf. Ser., vol. 1004, no. 1, 2018, doi: 10.1088/1742- 6596/1004/1/012029.
[13] Mauricio Menegaz, "Understanding YOLO – Hacker Noon," Hackernoon. 2018.
[14] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," 2016, doi: 10.1109/CVPR.2016.91.
[15] Wu, R.B. Research on Application of Intelligent Video Surveillance and Face Recognition Technology in Prison Security. China Security Technology and Application. 2019,6: 16-19.
[16] Tian, J.X., Liu, G.C., Gu, S.S., Ju, Z.J., Liu, J.G., Gu, D.D. Research and Challenge of Deep Learning Methods for Medical Image Analysis. ActaAutomatica Sinica,2018, 44: 401-424.
[17] Jiang, S.Z., Bai, X. Research status and development trend of industrial robot target recognition and intelligent detection technology. Guangxi Journal of Light Industry, 2020, 36: 65-66.
[18] Krizhevsky, A., Sutskever, I., Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems,2012, 25: 1097-1105.
[19] Russakovsky, O., Deng, J., Su, H., et al. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision,2015, 115: 211-252.
[20] Girshick, R., Donahue, J., Darrel, T.,Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In: Computer Vision and Pattern Recognition. Columbus.2014, pp. 580-587.
[21] He, K.M., Zhang, X.Y., Ren, S.Q., Sun, J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. IEEE Transactions on Pattern Analysis & Machine Intelligence,2015, 37: 1904-1916.
[22] Girshick, R. Fast R-CNN.In: Proceedings of the IEEE international conference on computer vision. Santiago.2015, pp. 1440-1448.
[23] Ren, S.Q., He, K.M., Girshick, R., Sun, J. Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. Montreal.2016, pp. 91-99.
[24] Redmon, J., Divvala, S., Grishick, R., Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In: Computer Vision and Pattern Recognition. Las Vegas.2016, pp. 779- 788.
[25] Pushkar Shukla, BeenaRautela and Ankush Mittal, "A Computer Vision Framework for Automatic Description of Indian Monuments", 2017 13th International Conference on Signal Image Technology and Internet- Based Systems(SITIS), Jaipur, India, ISBN (e): 978-1-5386-4283-2, December-2015.
[26] G.Sabarmathi , Dr.R.Chinnaiyan (2018), "Envisagation and Analysis of Mosquito Borne Fevers – A Health Monitoring System by Envisagative Computing using Big Data Analytics" in ICCBI 2018 – Springer on 19.12.2018 to 20.12.2018 (Recommended for Scopus Indexed Publication IEEE Xplore digital library)
[27] G.Sabarmathi , Dr.R.Chinnaiyan, Reliable Data Mining Tasks and Techniques for Industrial Applications, IAETSD JOURNAL FOR ADVANCED RESEARCH IN APPLIED SCIENCES, VOLUME 4, ISSUE 7, DEC/2017,PP- 138-142, ISSN NO: 2394-8442
[28] Dr. M. Thangamani, Jafar Ali Ibrahim, Information Technology E-Service Management System, International Scientific Global Journal in Engineering Science and Applied Research (ISGJESAR). Vol.1. Issue 4, pp. 13-18, 2017. http://isgjesar.com/Papers/Volume1,Issue4/paper2.pdf
[29] Ibrahim, Mr S. Jafar Ali, K. Singaraj, P. Jebaroopan, and S. A. Sheikfareed. "Android Based Robot for Industrial Application." International Journal of Engineering Research & Technology 3, no. 3 (2014).
[30] Ibrahim, S. Jafar Ali, and M. Thangamani. "Momentous Innovations in the Prospective Method of Drug Development." In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing, pp. 37-41. 2018.
[31] Ibrahim, S. Jafar Ali, and M. Thangamani. "Prediction of Novel Drugs and Diseases for Hepatocellular Carcinoma Based on Multi-Source Simulated Annealing Based Random Walk." Journal of medical systems 42, no. 10 (2018): 188. https://doi.org/10.1007/s10916-018-1038-y ISSN 1311-8080, https://acadpubl.eu/hub/2018-119-16/1/94.pdf
[32] Jafar Ali Ibrahim. S, Mohamed Affir. A "Effective Scheduling of Jobs Using Reallocation of Resources Along With Best Fit Strategy and Priority", International Journal of Science Engineering and Advanced Technology(IJSEAT) – ISSN No: 2321- 6905, Vol.2, Issue.2, Feb-2014, http://www.ijseat.com/index.php/ijseat/article/view/62
[33] M. Thangamani, and Jafar Ali Ibrahim. S, "Knowledge Exploration in Image Text Data using Data Hiding Scheme," Lecture Notes in

Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2018, 14-16 March, 2018, Hong Kong, pp352-357 http://www.iaeng.org/publication/IMECS2018/IMECS2018_pp352-357.pdf

[34] M. Thangamani, and Jafar Ali Ibrahim. S. "Knowledge Exploration in Image Text Data using Data Hiding Scheme." Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2018, 14-16 March, 2018, Hong Kong, pp352-357 http://www.iaeng.org/publication/IMECS2018/IMECS2018_pp352-357.pdf

[35] S. Jafar Ali Ibrahim and M. Thangamani. 2018. Momentous Innovations in the Prospective Method of Drug Development. In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing (DMIP '18). Association for Computing Machinery, New York, NY, USA, 37–41. https://doi.org/10.1145/3299852.3299854

[36] S. Jafar Ali Ibrahim and Thangamani, M "Proliferators and Inhibitors Of Hepatocellular Carcinoma", International Journal of Pure and Applied Mathematics (IJPAM) Special Issue of Mathematical Modelling of Engineering Problems Vol 119 Issue. 15. July 2018

[37] Thangamani, M., and S. Jafar Ali Ibrahim. "Ensemble Based Fuzzy with Particle Swarm Optimization Based Weighted Clustering (Efpso-Wc) and Gene Ontology for Microarray Gene Expression." In Proceedings of the 2018 International Conference on Digital Medicine and Image Processing, pp. 48-55. 2018. https://dl.acm.org/doi/abs/10.1145/3299852.3299866

[38] Testing" , IEEE International Conference on Intelligent Computing and Control Systems, ICICCS 2017, 512 – 517

[39] Dr.R.Chinnaiyan, Abishek Kumar(2017) ,"Construction of Estimated Level Based Balanced Binary Search Tree", 2017 IEEE International Conference on Electronics,Communication, and Aerospace Technology (ICECA 2017), 344 - 348, 978-1-5090-5686-6.

[40] R.Chinnaiyan, S.Somasundaram (2012) , Reliability Estimation Model for Software Components using CEP", International Journal of Mechanical and Industrial Engineering (IJMIE) , ISSN No.2231-6477, Volume-2, Issue-2, 2012, pp.89-93.

[41] R.Chinnaiyan, S. Somasundaram (2011) ,"An SMS based Failure Maintenance and Reliability Management of Component Based Software Systems", European Journal of Scientific Research, Vol. 59 Issue 1, 9/1/2011, pp.123 (cited in EBSCO, Impact Factor: 0.045)

[42] R.Chinnaiyan, S.Somasundaram(2011), "An Experimental Study on Reliability Estimation of GNU Compiler Components - A Review", International Journal of Computer Applications, Vol.25, No.3, July 2011, pp.13-16. (Impact Factor: 0.814)

[43] R.Chinnaiyan, S.Somasundaram(2010) "Evaluating the Reliability of Component Based Software Systems " ,International Journal of Quality and Reliability Management , Vol. 27, No. 1., pp. 78-88 (Impact Factor: 0.406)

[44] Dr.R.Chinnaiyan, Abishek Kumar(2017), Estimation of Optimal Path in Wireless Sensor Networks based on Adjancy List, 2017 IEEE International Conference on Telecommunication,Power Analysis and Computing Techniques (ICTPACT2017) ,6,7,8th April 2017,IEEE 978-1-5090-3381-2.