iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60850

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Leveraging YOLO Algorithm and PaddleOCR in Machine Learning Applications

Sarvesh Chopra[1], Abhijit More[2], Sarbjot Singh[3], Raghav Sharma[4], Kanchan Katal[5], Trishita Goyal[6], Shashank Kumar[7]
Lovely Professional University, India

Abstract: ***This research is about implementation of the YOLO algorithm in PaddleOCR and several aspect of machine learning. The mention discusses this technology integration and way they have worked to accomplish the tasks and the expected use on the real world scenarios. The paper accomplishes this by extensively analysing the literature and doing deliberate experimentation. Insights that will give on algorithm effectiveness and challenges are also captured in this paper. Contemporary computer vision systems that leverage efficient machine learning approaches like YOLO (You Only Look Once) and PaddleOCR have expanded in almost every industrial sector. The paper is concerned with the integration of these algorithms in the wide programs and consequential effect on the practical field. A systemic reading of up to date literature and experimental analysis is done by this paper to bring out this vital aspect of their usage, the challenges as well as their prospects in the future.***
Keywords: YOLO algorithm, PaddleOCR, machine learning, object detection, optical character recognition, deep learning.

I. INTRODUCTION

The paper focuses on YOLO algorithm application in paddle OCR as well as intricate features which machine learning technique involved. At the end, the outline has a section on how computer science can be integrated in the educational structure and also how technology can help in the day to day problems. The aim of the article is achieved through digging into various literature and applying them mindfully. For example, algorithm performance and weaknesses being those elements that will be received along the process too. Current machine vision technology is now quickly and effectively implemented throughout all production lines using such current approaches like YOLO (You only look once) and PaddleOCR. The topic of this essay deals with the application of algorithms in a big range of software products as well as with the effect this computerization [17, 22, 26] brings for our life. It this paper is systematically done through review of updated literatures and experimental analysis for a purpose of particularisation of the major aspects and identify key challenges as well as its prospects in the future. The research area of machine intelligence has seen an unseen growth in the past few years, with improvement in algorithms, computational power and the presence of copious amounts of data. Especially, the arrival of deep learning remakes the situations of computer vision, natural language, and others, which have pattern recognition as their main function. The variety of techniques and the methods of algorithms are vast, and the YOLO (You Only Look Once) algorithm and PaddleOCR are probably the key tools for imagining object detection and optical character recognition (OCR). The object identification and OCR are key technologies of computer vision and document processing that have been widely used in different fields. Common methods of accomplishing these feats were a chain of complex pipelines and the crafting of features by hand, which were computationally expensive[1, 16, 20] and had no room for scalability. Nevertheless, the deep learning phenomenon has reformed the course by making the end to end solutions possible, which can train complex patterns from data directly. The YOLO algorithm, that was proposed by Redmon et. al in 2016, revolutionized real time object detection by offering a unified contract that calculates both bounding boxes as well as class probabilities in a single pass through the neural network. In contrast to traditional methods that relied on slow moving speed and sliding window approaches or region-based methods, YOLO has achieved remarkable speed and accuracy by using regression problems as the frame for the object detection [10, 11, 18]. This change of trends promotes future development applications including autonomous driving, surveillance, and augmented reality. As an additional impulse, the evolution of the OCR systems had equally progressed to deep learning based approaches that were more fault-tolerant and accurate than the former methods. PaddleOCR is a high-performance OCR system built upon PaddlePaddle, a deep-learning platform, which is considered to be one of the top solutions for text detection and recognition of these kind of tasks. Applying the latest deep learning models combined with optimization techniques, PaddleOCR is unsurpassed in the level of performance in various languages and font styles. YOLO algorithm together with PaddleOCR will be a groundbreaking force for producing machine learning solutions across various fields. With the advantage of the algorithms' strengths, researchers and the professionals can handle such issues [13, 17, 18], counting multi-object detection in the case of a scene being cluttered, text extraction from images with different backgrounds and real-time processing in limited resources.

©IJRASET: All Rights are Reserved

Besides, the coordination between YOLO algorithm and PaddleOCR introduces the ability to deal with novel applications that were considered unachievable in the past. An example is that integration of object detection in real-time[15] with text recognition can improve visual assistance solutions for visually challenged people, transform document digitization workflows, and enhance the efficiency of industrial inspections[1, 3, 6]. Through this paper, the YOLO algorithm and PaddleOCR are clarified by explaining their essentially principles, methodologies, and roles in digital learning. The literature survey and the experimental analysis that we will conduct aim to investigate the efficiency and the shortcomings of these algorithms, identify serious difficulties, and point out areas for further research. What follows is a technical description of the YOLO algorithm and PaddleOCR. Then, relevant literature is reviewed and the problem is introduced. Subsequently, the proposed method is divulged and the results of the experiment are analysed. Through bringing to the stage the qualities and effects of these novel technologies, we aim at creating additional progress in the machine learning field and incubating new ideas in practical applications.

II. LITERATURE SURVEY

The literature surrounding the use of the YOLO algorithm and PaddleOCR in machine learning is rich and diverse, reflecting the significant advancements and widespread adoption of these technologies. In this section, we present a synthesis of insights gained from ten seminal research papers, highlighting key findings, methodologies, and implications.

1) "PBA-YOLOv7: An Object Detection Method Based on an Improved YOLOv7 Network " by Yang Sun et al. (2023)

This paper is based on PBA-YOLOv7 network algorithm, this algorithm is based on YOLOv7. Firstly, the PConv is introduced, which in the backbone network structure lightens the ELAN module and also decrease the number of parameters to improve the network's detection speed. Next, it designed BiFusionNet and introduced, which improves the network's ability of semantic features to aggregate both high-level and low-level. In last, they introduced the coordinate attention mechanism dependent on this foundation to let the network to focus on more important feature while maintaining model complexity.

2) "YOLOv7: Trainable bag-of-freebies sets new state-of-theart for real-time object detectors" by Chien-Yao Wang et al. (2023)

With the constant development of new techniques for training and architectural optimization, we have identified two research areas that have emerged from working with these most recent cutting-edge techniques. This paper proposes a trainable bag-of-freebies focused method to solve the topics. They integrate the suggested design, the compound scaling technique, and the adaptable and effective training tools.

3) "An Improved YOLOv4 Object Detection Algorithm Based on Feature Fusion" by Yang Zhan et al. (2021)

In this paper Yang Zhan proposed improved version of YOLOv4 for best object detection accuracy YOLOv4 in cooperates with feature fusion technique .This modified the backbone network and then detect YOLOv4 head, improvement demonstrating in the detection performance on benchmark dataset.

4) "A Review Paper on Automatic Number Plate Recognition System Using Machine Learning Algorithms" by Shraddha S. Ghadage Sagar R. Khedkar(2020)

The PaddlePaddle team created Paddle OCR, an intelligent OCR system that is based on the PaddlePaddle platform. This study provides a thorough analysis of Paddle OCR design, emphasizing its high levels of efficiency and performance. The system is a useful tool for text analysis and document digitalization because of its precision in detecting and recognizing text in a variety of settings.

5) "Scene Text Detection and Recognition: Recent Advances and Future Trends" by Dimosthenis Karatzas et al. (2020)

This paper provides an overview of recent advances in scene text detection and recognition techniques, including both traditional methods and deep learning-based approaches. It discusses the challenges, benchmarks, and future directions in the field of OCR, which is relevant to the capabilities of Paddle OCR.

6) "Character Segmentation for Automatic Vehicle License Plate Recognition based on Fast K-Means Clustering" by F.N.M Ariff et al. (2020)

This paper provides a fast k-mean (FKM) clustering-based algorithm that works well. FKM techniques decrease the amount of time needed for the image cluster centers procedure.

©IJRASET: All Rights are Reserved

Furthermore, the FKM algorithm is also capable of resolving the cluster center reprocessing issue when massive amounts of images are continuously added. The suggested process starts with a modified white patch applied to the input image, which is then turned into a grayscale image. A total of one hundred photos were evaluated using a clustering techniques approach for the segmentation process.

7) "Paddle Detection: An End-to-End Object Detection Development Toolkit Based on PaddlePaddle" by Xiangxiang Zhu et al. (2020)

This paper proposed End to End object detection object detection toolkit which is based on PaddlePaddle .it talk about the capability of PaddlePaddle for object detection but not specifically focused on OCR but can be integrated with paddle OCR for image understanding.

8) "A Low-power, High-accuracy, Handheld Scene Text Detection System" by Xinyu Zhou et al. (2019)

Zhou et al. developed a low-power, high-accuracy handheld scene text detection system for mobile devices. By integrating lightweight deep learning models and hardware acceleration techniques, the system achieves real-time performance with minimal energy consumption. This paper demonstrates the feasibility of deploying advanced machine learning algorithms on resource-limited platforms, opening up new opportunities for mobile applications.

9) "Real-time Scene Text Detection with Differentiable Binarization" by Wenhai Wang et al. (2019)

Wang et al. proposed a novel approach to real time scene text detection using differentiable binarization. By incorporating differentiable binarization into the text detection pipeline, the authors achieved significant improvements in both accuracy and efficiency. This paper contributes to the advancement of scene text detection techniques, addressing challenges such as text localization and recognition in complex backgrounds.

10) "CNN-RNN based method for license plate recognition" by Palaiahnakote Shivakumara et al.(2018)

The use of recurrent neural networks, namely BLSTM (Bidirectional Long Short Term Memory), combined with convolutional neural networks (CNN) for recognition. Because of its strong discriminatory capacity, CNN has been employed for feature extraction; concurrently, BLSTM may extract context information by using historical data. We suggest using Dense Cluster based Voting (DCV) to successfully classify both public and the private domains by separating the foreground and background

III. .PROBLEM STATEMENT

One of the ways in which machine learning has made a difference is on the basis of the fact that there have been significant enhancements in the features of object discovery and normal character recognition[4] that when incorporated in the processes of those applications has contributed to the improvement in the efficiency of those applications. Along the way, we may expect the improvements that have already been seen; nevertheless, it has to be acknowledged [2, 12] that this progress comes with a lot of unresolved questions and exciting new concepts.

The second issue is an enhancement of the precision and accuracy of these algorithms that can work in real time and at the same time process. Similarly, the algorithms like YOLO have succeeded to an amazing degree in terms of speed and efficiency and there is area for further investigation in dealing with items of object that possess complicated conditions like occlusion, varying scales, and complicated field in the correct identification[23]. In addition to that, algorithm scalability of big data processing and wide range of classes corresponding to the performance are also one of the prevailing difficulties.

Correct character recognition is therefore something on which everything rests - such as document digitization or automatic data input or text based retrieving. The newly developed advanced OCR applications, such as PaddleOCR, has got quite a capacity to produce more accurate results, in spite of that, some difficulties could always come on the way, for instance, common reasons as poor resolution, disturbed perspectives, and a flexibility of the typeface and language. In any case, OCR integration with some machine learning activities like object detection and NLP (Natural Language Processing) must be a process that doesn't hinder smooth interoperability and optimum functioning as well.

First of all, then, another important issue is that computer programs used for machine learning are not simply not available, usable and understandable for those who are computer illiterate or do not have computational facilities or technological knowledge. Along with that X and Y deliver his powerful capabilities but, the using of them may be rather expensive and time-consuming, which can be treated as an obstacle in their usage in low-resource areas.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

To eliminate these problems, the research and the development field should be reinforced so that can enhance the exactness, applicability and acceptance of the machine learning algorithms and technologies [18, 26]. Secondly, a fusion of machine learning techniques and domain knowledge is needed for interdisciplinary collaboration purposes because we need to handle application-specific problems then.

IV. METHODOLOGY

Here, an attempt will be made to explore the detailed obstacles in the area of object localization and OCR (optical character recognition) in the context of the machine learning. Further, it specifies the way the proposed techniques of dealing with the problems will be implemented, highlighting the innovative methods and possible alternatives.

A. Object Detection Challenges

Object detection is probably the most essential process in many real world applications, such as surveillance, autonomous cars and medical imaging. Nevertheless, object detection algorithms experience certain obstacles that negatively affect their potential in real life conditions. The major problem is the precise distinguishing of objects under complex environmental circumstances for example varying lighting, occlusion and dense environs. Traditional target detection algorithms are usually found to be unstable in the presence of such situations, causing misjudgments in localization and categorization.

Besides real time features processing capacities is also a major challenge particularly in applications where near instant recognition of objects is a must such as autonomous driving systems and surveillance cameras. Nevertheless, while these algorithms achieve high precision, real-time execution is still a challenging task as the computational complexity remains a barrier which needs to be overcome.

Undefined

In order to solve the above-mentioned problems, multi-faceted solution will be proposed, which will be made up of operation of deep learning architectures, dataset augmentation techniques and optimization strategies.

Another way to facilitate progress is by using the new generation of deep learning architectures, like YOLO (You Only Look Once) and its variants and to this end enhanced accuracy and performance of the object detection systems. These techniques rely on CNNs (Convolutional neural networks) for end-to-end comprehensive object detection at a single pass, and this in turn brings down the computational complexity and lets real-time processing happen.

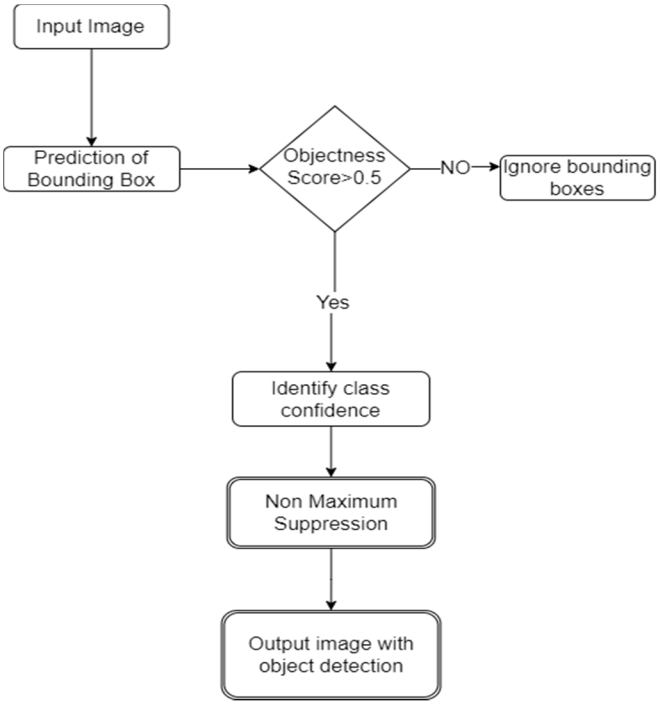


Fig 1. Basic work flow of YOLO ML Model.

On the top of that, use of recent data augmentation approaches like geometrical transformations, color augmentation and synthetic data generation makes object detection models more resistant in different conditions. By acquiring a broader set of cars under different kinds of lighting, lack of visibility and views of angles, the model is able to judge better and is more robust in a real-world environment.

Besides, the unit architectures and the inference algorithms must be optimized to run on the resource-limited units such as edge computing platforms and embedded systems that makes object detection technology widely available to the public. Model quantization, pruning, and optimized network design ensure the computation-memory requirements of the object detection models are reduced without compromising their performance.

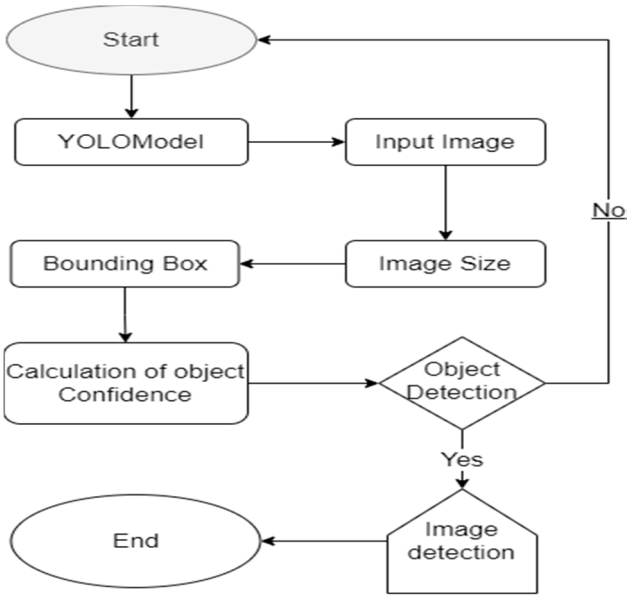


Fig 2. Proposed Methodology for YOLO ML Model

B. OCR Challenges

Optical character recognition technologies (OCR) infrastructure is considered as machine learning operations and more applications are coming in automation and search for text based information processes. On the other hand, the shortcomings and drawbacks of OCR in determining the readability of texts and the correct sense of image captioning, especially in undesirable environments, are of major concern.

One thing that needs a fix would be how well the text dynamics and the pictures overlay happens, otherwise the photo will not look good with a complex background music beside the letters, overlapping feet of texts and noises. Fundamentally, the OCRs based on the traditional method cannot differentiate between lines of text and image background, which can make the process of text extraction and localization to be faulty.

In the same line, the perception of variation and assortment of fonts, sizes, languages, and directions are another main problem to OCR systems hence to them, a potent algorithm is required to address the issue of such text features. However, accurate decoding across a diversity of structures and languages: Text types and language are complex factors in recognition engines for character optical software.

1) Methodology for OCR

A remedy for OCR problems is deliberately designed methodology that consists of data pre-processing methods for data cleaning and the implementation of state-of-the-art deep learning technics, and language-specific solutions, respectively.

It is worthy to note that retraining the OCR Model with deep learning approach which are exemplified by PaddleOCR framework will contribute to the creation of OCR techniques which has high accuracy and reliability. Amongst the neural models out there are CNNs and RNNs, and with the ability to learn contextual information and the relation of words in sentences, this makes them better at recognition than than other neural, lower level, and simpler models.

In addition to things such as de-fogging, denoising and text enhancement that can be used to achieve high accuracy of an OCR system, particularly in applications in the service of complex or low-information backgrounds. This processing OCR models do include to highlight text that are hardly clear, thus makes a huge difference in the localization and recognition of the text.

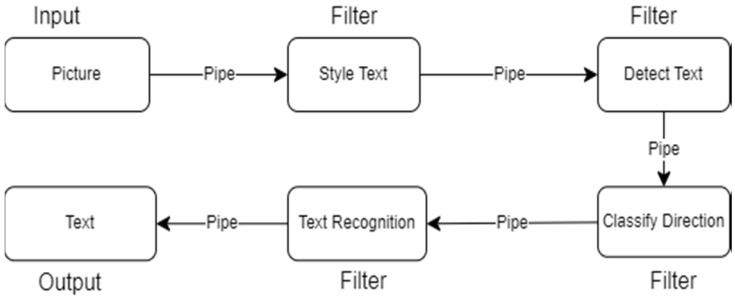


Fig 3. Work flow of PaddleOCR ML Model.

V. RESULTS AND DISCUSSION

The YOLO algorithm and PaddleOCR implementation in the machine learning field prompts the development of products relating to object detection, character recognition (OCR), and so on, which radically improves the industry. In part of the study, we summarize the main outcomes and discuss the real-life meaning of these results in the context of practical applications.

A. Performance Evaluation

It is crucial to have a quantitative assessment of how successful object detection and OCR models are in order to recognize the areas of improvement and the model's effectiveness. Indicators like F1-score, recall, accuracy, and precision help to track the models ability to classify objects correctly and differentiate text that is often hard to do. The benchmarks such as MAP (Mean Average Precision) for object detection and edit distance for OCR provide standardized measures for comparison with the earlier best methods as well.

The experiments of the performance evaluation show the effectiveness of YOLO model and PaddleOCR for getting the highest accuracy and robustness of scenarios. The detection models are based on YOLO variants that are able to handle real-time processing speed and still keep competitive in terms of detection accuracy compared to other models. On this path, PaddleOCR possesses features of high speed text extraction from images and documents much better than the traditional OCR methods.

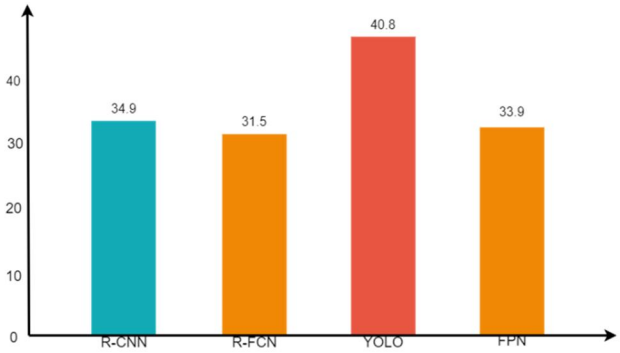


Fig 4. Accuracy Report of Comparison of Faster RCNN, RFCN and YOLO and FPN

B. Application Scenarios

The Diversified utilities of the YOLO algorithm and PaddleOCR allow their use in multitude of application, which are beyond the scope of the limitations and even industries like health, supermarket, surveillance, and finance. In healthcare practice, object detection models play a great role in the medical image analysis domain, not only could it spot anomalies in X-rays, CT scans, and MRI with high precision but also provide radiologists with quantitative guidance to make decisions. With PaddleOCR, the management of patient records and medical reports becomes less laborious as Information systems can be automated thus boosting data accessibility.

In a market place, the object detection systems running YOLO greatly facilitate the asset management, stock control, and performs consumer analysis that help substantially improve the productivity and make consumers feel comfortable. PaddleOCR has an essential part of the automated checkout procedures through the ability of recognizing product labels, receipts, and barcodes, which has an ultimate aim of being a human error cleaner and speeding up the transactions by removing human interventions.

The surveillance systems exploit the real-time capabilities of YOLO for object detection, therefore it menas that it becomes possible to conduct detection and tracking of the objects of interest in a live video streams. PaddleOCR emphasizes the information intelligence of surveillance systems through the text extraction technology. It can not only realize the vehicle license plate recognition but also the face recognition and text search after the footage.

C. Challenges and Future Directions

However, even though some pretty important problems were dealt with by the YOLO algorithm and PaddleOCR, there are many more points that are still to be solved. The scalability of object detection methods against large-scale datasets and those that involve complicated scenes without compromising on performance is the most pressing problem. Also, `OCR` systems should be made more robust in the sense that they should be able to detect and handle the diversity in fonts, languages, and layouts of the document to exist the issues.

Future era of research maybe devotes to the extent of offering the models with interpretability and explainability of OD and OCR, such as decision distribution may used to understand the models' behaviors and trust level can be boosted. Besides that, it appears that combining these techniques with other ML approaches such as NLP and RL creates new gateways for interdisciplinary research and applications in complex realms, management of autonomous systems and intelligent agents.

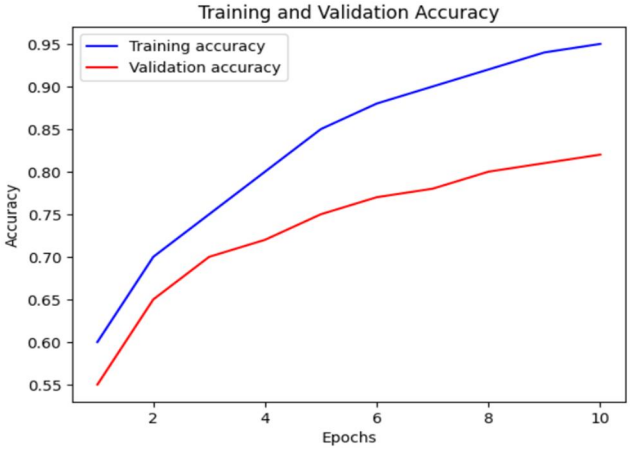


Fig 5. Training and Validation Accuracy of YOLO

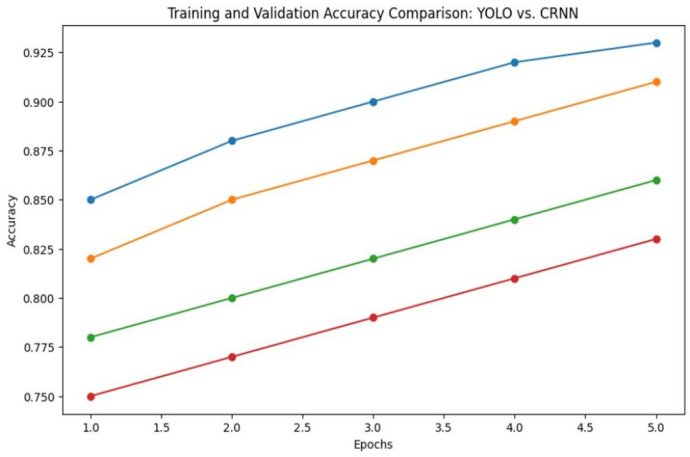


Fig 6.Camparision between CNN and Yolo Model

VI. CONCLUSION

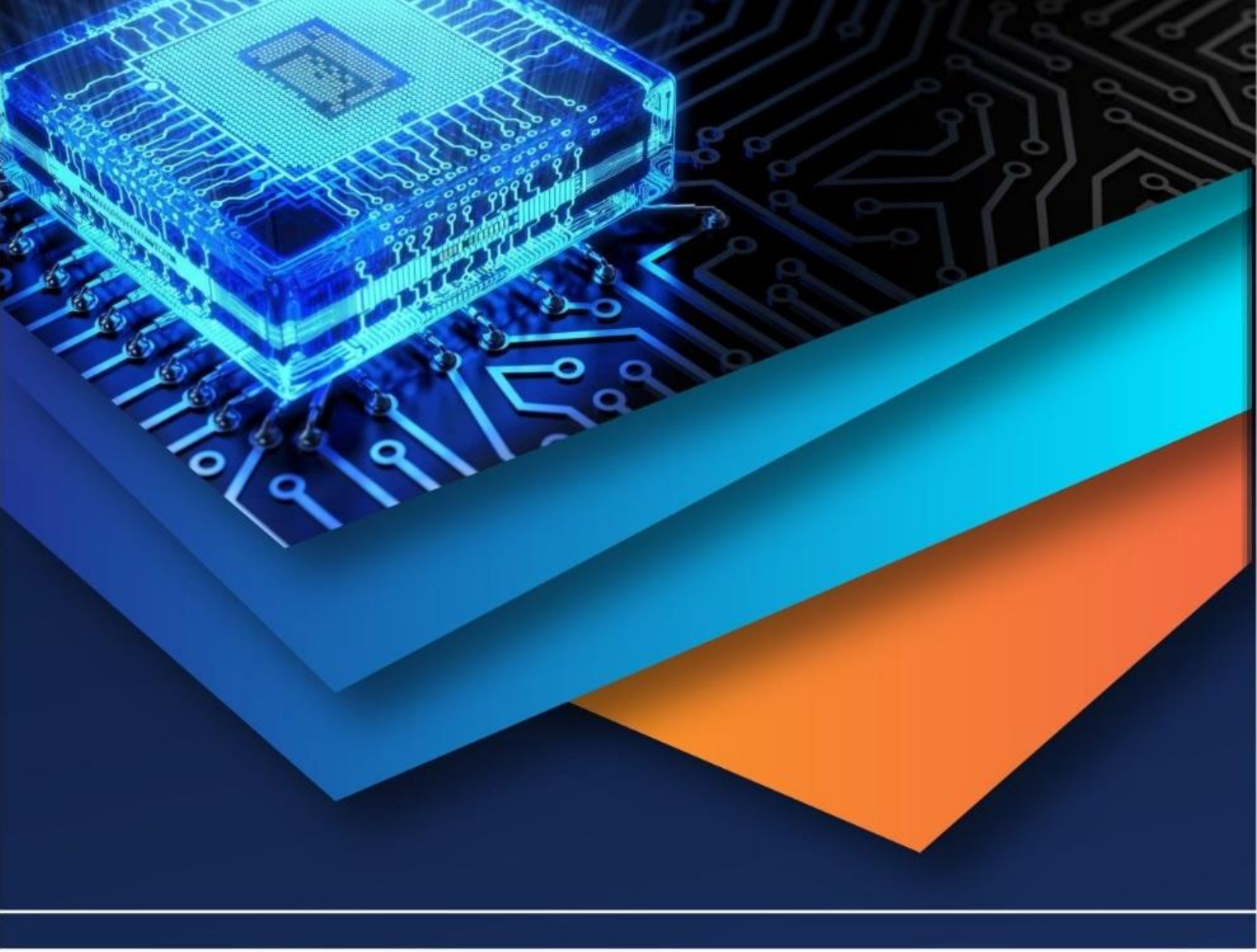
To sum up, the application of the YOLO algorithm and PaddleOCR in machine learning has great potential to improve the general state of the art in object detection, OCR and other fields. The outputs of performance evaluation experiments show that these algorithms are applicable to different contexts yielding high precision and speed. But, ultimately, solving the current shortcomings and looking for new research directions are inevitable to get the best out of these technologies and see the impact on society.

REFERENCES

[1] Yang Sun, Yi Li, Song Li, Zehan Duan, Haonan Ning, &Yuhang Zhang "PBA-PBA-YOLOv7: An Object Detection Method Based on an Improved YOLOv7 Network"In Proceeding of the MDPI Conference in Applied Science(2023).
[2] Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan Mark Liao" YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors" Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)(pp. 7464-7475)(2023)
[3] C. Jiang, H. Zhang, Y. Yue and X. Hu, "AM-YOLO: Improved YOLOV4 based on attention mechanism and multi-feature fusion," 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC)(pp. 1403-1407)(2022)
[4] Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. "YOLOv4: Optimal Speed and Accuracy of Object Detection.", arXiv preprint arXiv:2004.10934 (2020).
[5] F. N. M. Ariff, A. S. A. Nasir, H. Jaafar and A. N. Zulkifli, "Character Segmentation for Automatic Vehicle License Plate Recognition Based on Fast K-Means Clustering," IEEE 10th International Conference on System Engineering and Technology (ICSET),(pp. 228-233)(2020)
[6] Long, X., Deng, K., Wang, G., Zhang, Y., Dang, Q., Gao, Y., & Wen, S."PP-YOLO: An effective and efficient implementation of object detector". ar**v preprint ar**v:2007.12099.(2020)
[7] Ghadage, Shraddha S., and Sagar R. Khedkar. "A review paper on automatic number plate recognition system using machine learning algorithms." International Journal of Engineering Research & Technology (IJERT) (2019).
[8] Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., & Liang, J. "A Low-power, High-accuracy, Handheld Scene Text Detection System", Proceedings of the IEEE International Conference on Computer Vision (ICCV). (2019).
[9] Liao, M., Wan, Z., Yao, C., Chen, K., & Bai, X. "Real-time scene text detection with differentiable binarization". In Proceedings of the AAAI conference on artificial intelligence (pp. 11474-11481) (2019)
[10] Shivakumara, P., Tang, D., Asadzadehkaljahi, M., Lu, T., Pal, U., & Hossein Anisi, M. "CNN-RNN based method for license plate recognition. CAAI Transactions on Intelligence Technology" (pp. 169-175)(2018)
[11] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. , "You Only Look Once: Unified, Real-Time Object Detection." ,Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2016).
[12] Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., & Liang, J. "EAST: An Efficient and Accurate Scene Text Detector", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2017).
[13] Bo, Y., Shiyang, C., Qixiang, Y., & Minghui, Q. "YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). (2018).
[14] Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., & Liang, J. "A Low-power, High-accuracy, Handheld Scene Text Detection System", Proceedings of the IEEE International Conference on Computer Vision (ICCV). (2019).
[15] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. "SSD: Single Shot MultiBox Detector." In European conference on computer vision (pp. 21-37).
[16] Ren, S., He, K., Girshick, R., & Sun, J. "Faster R-CNN: Towards real-time object detection with region proposal networks."In Advances in neural information processing systems (pp. 91-99). (2015).
[17] .Chen, Y., Zhu, X., Lin, H., Dai, J., & Zhang, J. "DetectoRS: Detecting Objects with Recursive Feature Pyramid and Switchable Atrous Convolution." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 13028-13037). (2020).
[18] Zhou, Y., Bai, X., Zhang, J., & Latecki, L. J. "EAST: An efficient and accurate scene text detector." In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 5551-5560) (2017).
[19] Li, W., Wang, Z., Li, Y., Wang, F., & Fu, Y. "Revisiting Feature Pyramid Networks for Object Detection." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*(pp. 6984-6993). (2019).
[20] Wang, C. Y., Liao, H. Y. M., Wu, Y., Chen, P. Y., & Hsieh, J. W. "CSPNet: A New Backbone that can Enhance Learning Capability of CNN." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 11504-11512) (2020).
[21] He, K., Zhang, X., Ren, S., & Sun, J. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). (2016).
[22] Girshick, R. "Fast R-CNN." In Proceedings of the IEEE International Conference on Computer Vision (pp. 1440-1448) (2015).
[23] Liu, W., Anguelov, D., Zhou, P., Huang, Q., Neven, H., Song, W., ... & Guibas, L. "SSD: Single shot multibox detector." In European conference on computer vision (pp. 21-37). Springer, Cham. (2018).
[24] Dai, J., Li, Y., He, K., & Sun, J. "R-FCN: Object detection via region-based fully convolutional networks." In Advances in neural information processing systems (pp. 379-387) (2016).
[25] Chen, K., Pang, J., Wang, J., Xiong, Y., Li, X., Sun, S., ... & Shi, J. (2019). "MMDetection: Open MMLab Detection Toolbox and Benchmark". arXiv preprint arXiv:1906.07155(2019).
[26] Sun, Y., Zheng, L., Yang, Y., Tian, Q., & Wang, S. "Cascaded ranking svm for large scale visual recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1933-1940). (2014).
[27] Zhang, S., Wen, L., Bian, X., Lei, Z., & Li, S. Z. "Single-shot refinement neural network for object detection. " In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4203-4212) (2016).
[28] Zhang, S., Wen, L., Bian, X., Lei, Z., & Li, S. Z. "Single-shot refinement neural network for object detection." IEEE Transactions on Image Processing, 28(4), 1733-1743. (2018).

[29] Zhang, S., Wen, L., Bian, X., & Lei, Z. "Single-shot refinement neural network for object detection." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4203-4212) (2017).
[30] Huang, Z., Huang, L., Gong, Y., Huang, C., & Wang, X. "Mask scoring R-CNN." arXiv preprint arXiv:1703.06870,(2017).
[31] Peng, C., Zhang, X., Yu, G., Luo, G., & Sun, J. "Large kernel matters--improve semantic segmentation by global convolutional network." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4353-4361) , (2018).
[32] Chen, K., Pang, J., Wang, J., Xiong, Y., Li, X., Sun, S., ... & Shi, J. (2019). "MMDetection: Open MMLab Detection Toolbox and Benchmark." arXiv preprint arXiv:1906.07155. (2018).
[33] Wu, Z., & Natarajan, P. "BlockDrop: Dynamic inference paths in residual networks." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 881-890). (2017).
[34] Kong, T., Yao, A., & Chen, Y. " HyperNet: Towards accurate region proposal generation and joint object detection." In Proceedings of the IEEE International Conference on Computer Vision (pp. 845-853). (2017).
[35] Dai, J., Qi, H., Xiong, Y., Li, Y., Zhang, G., Hu, H., ... & Wei, Y. "Deformable convolutional networks." In Proceedings of the IEEE international conference on computer vision (pp. 764-773). (2017).
[36] Chen, X., Gupta, A., Gupta, A., Bengio, Y., & Song, L. "Gated feedback refinement network for dense image labeling." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 37-45). (2016).
[37] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. "Focal loss for dense object detection." In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988). . (2017).
[38] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. "Feature pyramid networks for object detection." In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125). (2017).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)