International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Image Recognition Using Machine Learning and Yolo Algorithm

Dr.R.Nagalakshmi B.E., M.E.,Ph.D Assistant Professor.
Department of CSE
SRM IST
Ramapuram,Chennai,India
nagalakr1@srmist.edu.in

V.Tej Manendra
Department of CSE
SRM IST
Ramapuram,Chennai,India
vm1473@srmist.edu.in

B.Nikhil Srinivas
Department of CSE
SRM IST
Ramapuram,Chennai,India
bs6186@srmist.edu.in

CH.Ayyappa
Department of CSE
SRM IST
Ramapuram,Chennai,India
cc2160@srmist.edu.in

***Abstract*— Since olden times, the identification has remained a core topic of study. Research progressed tremendously after the development of deep learning models, which greatly enhanced the precision of representation classification. Two models were developed for image recognition in this research paper. Additionally, two models were constructed using deep learning techniques: a deep face network and an encapsulating network. Besides, this paper focuses on various face recognition applications which include: object labeling, first recognition, and therapeutic concept analysis. This paper carries out an extensive work on the current matters regarding the figure recognition in machine learning and elaborates on the primary challenges and future possible areas of research. Attempts of deep learning systems with the use of artificial neural networks aim to emulate human intelligence processes. Real-time image recognition entails the instant tagging of objects, yielding precise and efficient outcomes for specific actions or events. The spectrum of computer science vision encompasses a wide range of machine vision problems. The use of radio signals for various purposes had inspired some researchers to come up with more precise and sophisticated predictions.**

I. INTRODUCTION

Image acknowledgment merges a range of algorithms to produce machine representations from images, which are grouped according to their similarities. The various classifications captured during image classification include delineation of the primary focus captured within an image. Object localization entails marking one or more regions in a given image, and placing a bounding box around the marked area. Object identification integrates the processes of marking and monitoring multiple objects in an image, and issuing a classification to a given area within the image. Often, people talking about object recognition actually mean detecting the Concept relegation allows a class name to function as a prototype while object localization requires the movement of a bounding box over one or more targets in an image. Detecting objects is now more sophisticated having two functions - identifying the objects and classifying them in an image. Such problems are usually associated with object recognition, which is very much associated with other methods in computer vision such as image detection and image segmentation. Object recognition is of utmost importance for understanding and interpreting visual information contained in images or videos. Object localization is useful in many research cases for purposes ranging from protecting the individual to increasing productivity at the workplace.

Facial recognition can be useful for security purposes by allowing only certain people access to particular areas within a system. It can also be used to monitor the number of people within a facility, thus improving productivity for that group. Moreover, it is utilized in visual searching systems so that users can easily search for certain items. The value of this technology is apparent in social and e-commerce based platforms where it enhances numerous services. These functions take advantage of the classifying of people and things as a means for organizing and controlling data in professional settings. Image classification is concerned with the predicting the area of interest in an image and hence assigning a class to that image.

II. LITERATURE REVIEW

Ms. Sharma et al. [1] (2021) conduct a comprehensive examination of object detection through the use of OpenCV and Python, with particular emphasis on real-time applications such as automotive safety systems and pedestrian monitoring. They highlight the adaptability of OpenCV, which accommodates a range of machine learning algorithms. The study includes practical applications, such as Haar Cascade classifiers and frame differencing, illustrating their significance in intelligent vehicle systems. Furthermore, the authors present a project tailored for beginners, outlining practical steps for implementing object detection. This paper evaluates both traditional object, so in terms of people using the definition that is how most people understand it. It is enjoyable for learners to make up or try to distinguish other similar tasks under computer vision. OpenCV is a complete collection of algorithms for computer vision, alongside other functions necessary for the implementation of OpenCV algorithms. Open CV commands special computer vision functions accessible for free using Python, C++, Java and many more.

Kavya Gupta [2](2023) examines the progress made in machine learning methodologies pertaining to object detection and classification, emphasizing the challenges and innovations present in multimedia and intelligent transportation systems. The review underscores the significance of Convolutional Neural Networks (CNNs) and sophisticated feature extraction techniques in enhancing both detection accuracy and processing speed. Gupta discusses challenges such as object diversity and the importance of temporal context in video analysis, while also investigating the utilization of these techniques in remote sensing, particularly through hyperspectral imaging for target detection and land cover classification.

The research by Shridevi Soma and Nischita Waddenkery [3] focuses on developing an advanced video surveillance system using the YOLOv3 object detection algorithm, known for its real-time precision. The study aims to detect and classify multiple objects, such as individuals and vehicles, within a single frame, using an enhanced KITTI dataset that includes both foreground and background objects. YOLOv3's CNN-based approach predicts bounding boxes and class probabilities simultaneously, improving object localization through Non-Maximum Suppression (NMS).

The article entitled "Enhancing Object Recognition in Crime Scenes through Local Interpretable Model-Agnostic Explanations," authored by Helia Farhood, Morteza Saberi, and Mohammad Najafi and published in [4] 2021, tackles the issue of interpreting deep learning models utilized for object recognition in crime scene investigations. The authors emphasize that the intricate nature of these models renders their decision-making processes unclear, thereby hindering their effectiveness in forensic applications.

The document authored in 2021 by Md. Moshiur Rahman, Shajeeb Chakma, Sadia Akter, Dewan Mamun Raza, and Abdus Sattar [5] delves into the realm of real-time object detection, emphasizing the difficulties associated with the identification and localization of objects within images and videos across diverse conditions. This paper evaluates both traditional and machine learning-based detection methodologies, underscoring significant progress in deep learning, particularly through the implementation of the SSD (Single Shot MultiBox Detector) utilizing MobileNet V3. The model proposed exhibits proficiency in real-time detection.

The 2021 research by Ahmed Sheikh Abdullahi Madey, Amani Yahyaoui, and Jawad Rasheed [7]compares traditional and deep learning techniques for vehicle detection in videos. Presented at the International Conference on and machine learning-based detection techniques, underscoring significant progress in deep learning algorithms.

III. PROPOSED METHODOLOGY

A.Deep Reinforcement Learning

The submitted methods utilizes modern deep support learning (drl) methods to raise the effectiveness and elasticity of representation recognition structures. Drl, an creative approach that combines the substances of deep knowledge and reinforcement knowledge, admits models to acquire progressive ocular skills through interplay accompanying their surroundings. In this foundation, the model is deliberate an agent that interprets recommendation representations and learns to identify or classification objects by optimizing a reward signal established accurate prophecies. In contrast to unoriginal supervised education forms, drl enables bureaucracy to reinforce allure recognition competencies usually by learning from a succession of knowledge.

Fig.1 Statistical Data presentation for Reinforcement Learning

In the rule of concept recognition, drl allows the power to focus on main domains within an countenance, easing consideration-driven feature distillation. This procedure can be regulated in evident-time when confronting new deformities or patterns. This leads to upgraded accuracy in complicated positions where occlusions, environment crash, and mixed illumination survive. By including drl into image acknowledgment, bureaucracy can not only effectively categorize objects but more actively determine, steadily broadening its understanding of new object types and tangible environments. This plan advances the capabilities of wise optical systems, admitting ruling class to address a variety of palpable-globe challenges, containing self-driving machines and healing image reasoning. The submitted method exploits deep support education (drl) to tackle image acknowledgment issues and presents an common and adaptable ocular understanding method. To start, a diverse array of concept classifications is assembled and prepared utilizing methods such as resizing, normalization, and improving. These processes consist of conduct like turn, throwing, and scaling to advance the model’s ability to generalize. In this circumstances, the representation recognition plan is depicted as an power that engages accompanying allure surroundings, considering each concept as a state and performing conduct to classification them. Choosing the right determinations results in definite consequences, while mistakes lead to negative results, inciting the agent to polish allure strategies through continuous interplays accompanying the environment. Reinforcement education (rl) used to image acknowledgment determines a unique procedure for forwarding complex optical tasks by training an power to form decisions by way of a experimental approach process.

Forthcoming Networks and Sustainability in the AIoT Era, the study evaluates the performance of HOG with a Linear SVM classifier against the YOLOv3 deep learning algorithm. Using the KITTI and GTI vehicle image datasets, the research demonstrates that YOLOv3 offers superior real-time detection accuracy, despite challenges in shadowed or similarly colored environments. This study highlights the advancements in detection accuracy and speed through deep learning.

B.YOLO-Based Image Recognition

In the rule of concept recognition, drl allows the power to focus on main domains within an countenance, easing consideration-driven feature distillation. This procedure can be regulated in evident-time when confronting new deformities or patterns. This leads to upgraded accuracy in complicated positions where occlusions, environment crash, and mixed illumination survive. By including drl into image acknowledgment, bureaucracy can not only effectively categorize objects but more actively determine, steadily broadening its understanding of new object types and tangible environments. This plan advances the capabilities of wise optical systems, admitting ruling class to address a variety of palpable-globe challenges, containing self-driving machines and healing image reasoning. The submitted method exploits deep support education (drl) to tackle image acknowledgment issues and presents an common and adaptable ocular understanding method. To start, a diverse array of concept classifications is assembled and prepared utilizing methods such as resizing, normalization, and improving. These processes consist of conduct like turn, throwing, and scaling to advance the model's ability to generalize. In this circumstances, the representation recognition plan is depicted as an power that engages accompanying allure surroundings, considering each concept as a state and performing conduct to classification them. Choosing the right determinations results in definite consequences, while mistakes lead to negative results.

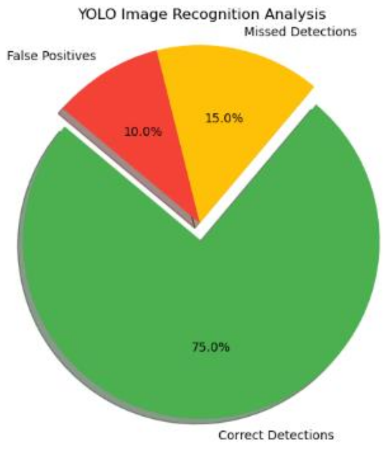

Fig.2 Analysis using YOLO

In exercise, the network gains information to reduce prognosis wrong by altering the model weights during the whole of training. After preparation the yolo model, it is proven accompanying confirmation data to judge allure veracity concerning precision, recall, and mean average accuracy (picture). When the model has proved acceptable acting, it is organized into physical-period object detection arrangements.

IV. RESULT AND DISCUSSION

The representation acknowledgment system grown in this place research was judged using a dataset that contained various objects. After extensive preparation and cautious adaptations, the model demonstrated superior veracity in understanding objects and processing dossier at an unusually hasty rate. The findings recorded that bureaucracy surpassed in object recognition inside elaborate environments and troublesome illumination positions. The average detection assurance usually waited over 90%, with only a restricted number of dishonest a still picture taken with a camera and missed detections, emphasize the dependability of the yolo foundation. Moreover, the system had the skill to process dossier in real-opportunity, making it acceptable for requests that required prompt accountable, to a degree self-forceful cars and safety schemes. In comparison to usual detection designs, yolo brought active results without negotiating veracity. Furthermore, a comprehensive reasoning of the yield figures showcased the model's talent to correctly recognize and classify diversified objects together. The analysis points out that while yolo acts unusually in most cases, challenges happen when difficult to discover very small objects or parts that are completely unseen, suggesting the necessity for further augmentations of the model or the unification of composite methods.

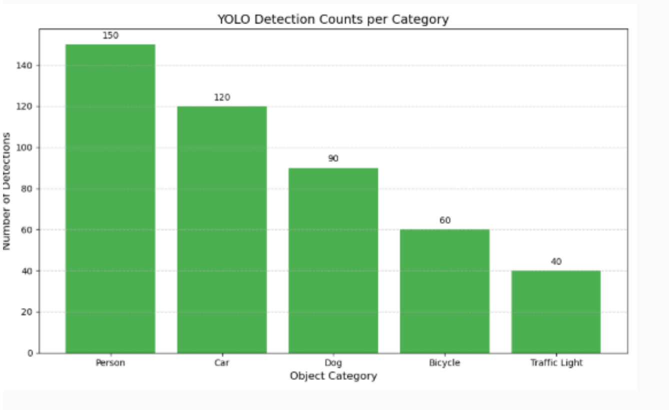

Fig.4 Final Analysis Part

V.CONCLUSION

In end, countenance recognition engaging YOLO (you only look formerly) is a reliable and persuasive technique for detecting objects in legitimate-period. YOLO uses an progressive deep education model that can swiftly resolve representations in a sole pass, greatly embellishing the discovery speed distinguished to common methods. By partitioning the concept into a gridiron, YOLO together predicts the object classes inside each division and their accompanying restricting boxes, which form it appropriate for uses that demand fast and correct outcomes. It outperforms different methods in conditions of efficiency and veracity, permissive the labeling of objects in an countenance with unusual accuracy. This facility is specifically advantageous for uses to a degree television surveillance, self-forceful cabs, and authentic-period image study. However, the depiction of YOLO is liable to be subjected the kind of the trained model and the dataset, alongside the computational money handy for handle. Overall, YOLO has emerged as a popular alternative for requests that make necessary fast and accurate object discovery across miscellaneous rules.

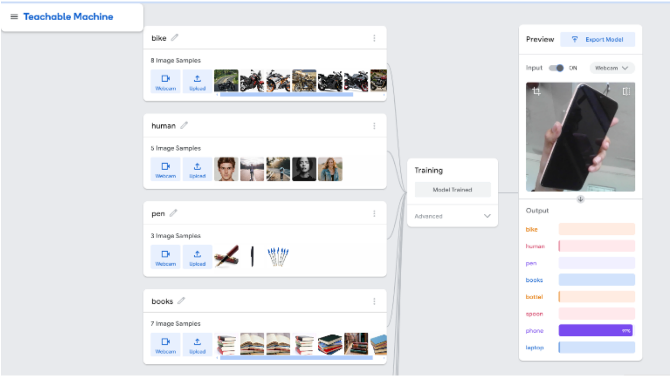

Fig.3 Image Recognition Training Analysis

REFERENCES

[1] Geethapriya S, N. Duraimurugan, S.P. Chokkalingam, "Real-Time Object Detection

with Yolo", International Journal of Engineering and Advanced Technology (IJEAT)

[2] Abdul Vahab, Maruti S Naik, Prasanna G Raikar an Prasad S R4, "Applications of Object Detection System",International Research Journal of Engineering and Technology (IRJET)

[3] Hammad Naeem, Jawad Ahmad and Muhammad Tayyab, "Real-Time Object Detection and Tracking",IEEE

[4] Meera M K, & Shajee Mohan B S. 2016, "Object recognition in images", international Conference on Information

[5] Viola, Paul & Jones, Michael. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. IEEE Conf Comput Vis Pattern Recognit. 1. I-511. 10.1109/CVPR.2001.990517.

[6] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... &Zitnick, C.L. (2014, September). Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755). Springer, Cham.

[7] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.

[8] Patterson, G., & Hays, J. (2016, October). Coco attributes: Attributes for people, animals, and objects. In European Conference on Computer Vision (pp. 85-100). Springer, Cham.

[9] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

[10] A. Heredia and G. Barros-Gavilanes, "Video processing inside embedded devices using SSD-Mobilenet to count mobility actors," 2019 IEEE Colombian Conference on Applications in Computational Intelligence (ColCACI), Barranquilla,Colombia,2019,pp.16,doi:10.1109/ColCACI.2019.8781798., Y. A. et al. "Enhanced Fraud Detection in UPI Transactions using Machine Learning Techniques." International Journal of Computer Applications 182.36 (2018): 45-51.