ISSN No:-2456-2165

Enhancing Automobile Manufacturing Efficiency using Machine Learning: Sequence Tracking and Clamping Monitoring with Machine Learning Video Analytics and Laser Light Alert System

Amrathakara Bhat[1]
[1]Research Associate, Innodatatics Inc., Hyderabad, India

Aishwarya Dhadd[2]
[2]Mentor, Research and Development, Innodatatics Inc., Hyderabad, India

Bhushan Sharad Patil[3]
[3]Team Leader, Research and Development, Innodatatics Inc., Hyderabad, India

Bharani Kumar Depuru[4*]
[4]Director, Innodatatics Inc., Hyderabad, India

Corresponding Author:- Bharani Kumar Depuru[4*]
OCR ID: 0009-0003-4338-8914

Abstract:- The research focuses on leveraging video analytics to track the sequence followed during clamping in the Automobile Manufacturing industry while adhering to the principles of the Poka-Yoke system. The study proposes the development of a Machine Learning (ML) model using the latest version of YOLOv8 (You Only Look Once) / YOLO-NAS (Neural Architecture Search) to achieve a remarkable accuracy rate of 99%. By harnessing the power of video analytics, the manufacturing process can be monitored and optimized to ensure efficient clamping operations. The utilization of video analytics enables real-time tracking of the clamping sequence, providing valuable insights into the production line. The ML model developed with YOLOv8 can accurately identify and analyze the clamping steps, ensuring that they followed the correct sequence. By adhering to the principles of the Poka-Yoke system, which is an error-proofing method, the manufacturing industry can significantly reduce defects and improve overall quality. The proposed system's integration with video analytics and ML techniques offers many advantages, including continuous monitoring, rapid identification of deviations, and immediate corrective actions. By achieving a 99% accuracy rate, the system provides a robust and reliable solution for ensuring precise adherence to the clamping sequence, contributing to enhanced manufacturing efficiency. The research also explores the potential deployment of the system in an AWS (Amazon Web Services) cloud environment, which offers scalability, flexibility, and efficient data processing capabilities. This cloud-based implementation allows for seamless integration into existing manufacturing workflows and facilitates centralized monitoring and management. Overall, this study presents a comprehensive approach to tracking the clamping sequence in the Automobile Manufacturing industry, leveraging video analytics, and adhering to the Poka-Yoke system. The ML model developed using YOLOv8 (You Only Look Once) / YOLO-NAS (Neural Architecture Search) demonstrates exceptional accuracy, paving the way for improved quality control, reduced errors, and enhanced productivity in automotive manufacturing processes.

***Keywords**:- Poka-Yoke, Automobile Manufacturing, Video Analytics, Artificial Intelligence, Total Quality Management in Manufacturing.*

I. INTRODUCTION

The aim of this research study and development of machine learning (ML) is to elevate the efficiency of automobile manufacturing. This research study and development is based on the open-source CRISP-ML(Q) mindmap available on the 360DigiTMG website (ak.1) [Fig.1]. In the manufacturing industry, manual assembly plays a crucial role in the production of various vehicle body parts, including truck bodies joining with precision. Manual truck body manufacturing involves assembling different vehicle parts to create a functional and durable truck body. However, this process can be prone to certain quality issues that need to be carefully addressed to ensure the production of high-quality truck bodies. Some issues common are Inconsistent Fit and Finish, Poor Weld Quality, Loose or improper joined parts, Inadequate Surface Preparation and coating because of gaps, and incomplete or Incorrect Component joining.

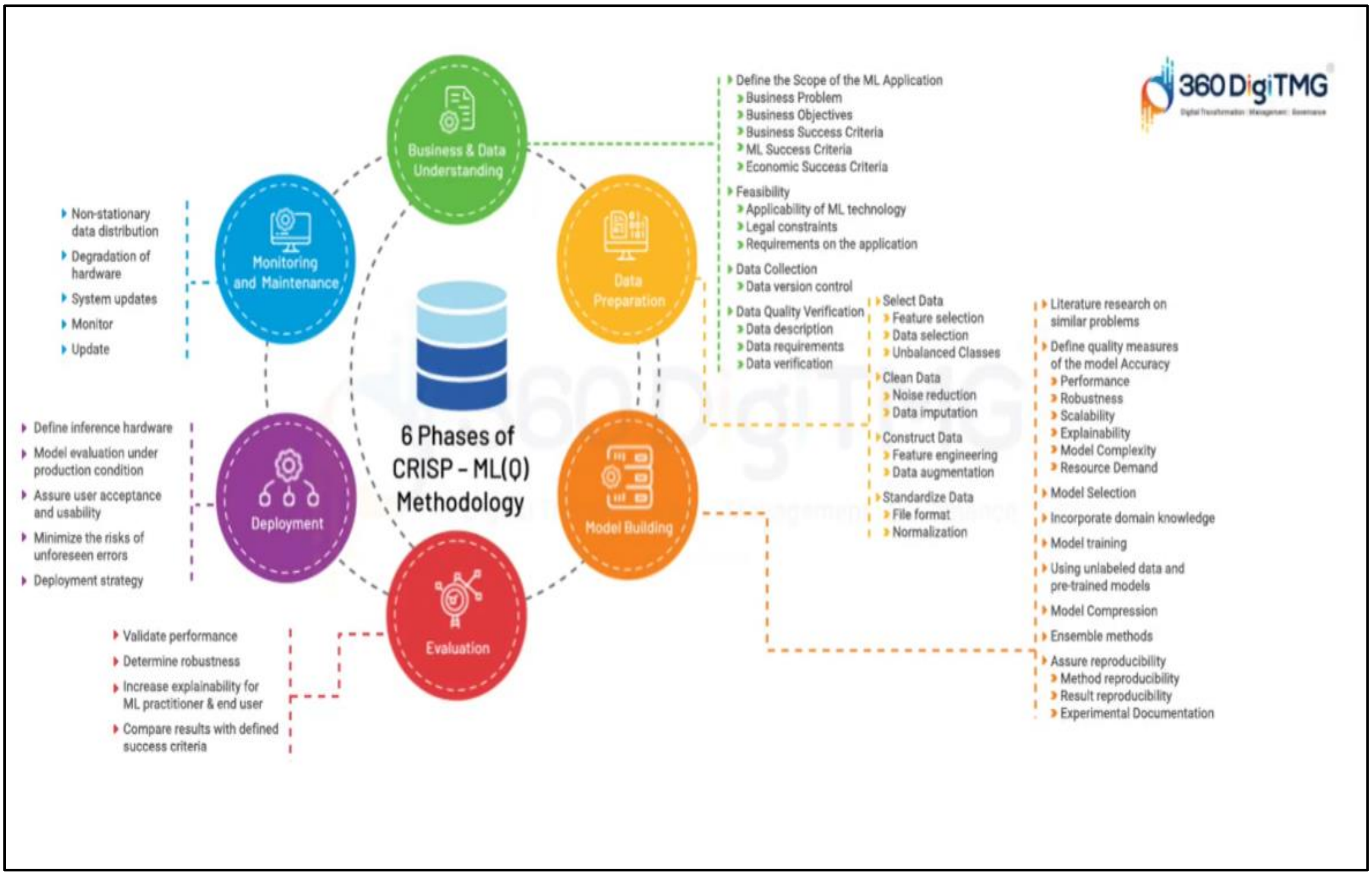


Fig 1 CRISP-ML (Q) Methodological Framework, Outlining its Key Components and Steps Visually.
(Source: -Mind Map - 360DigiTMG)

It is essential to address the human error of implementing the Poka-Yoke [20] system that may arise during the assembly process. By implementing proper training, quality control measures, and regular inspections, manufacturers can improve the consistency, precision, and overall quality of manual truck body assembly, ensuring the timely production of high-quality and reliable products. But practically this process is not followed because of human psychology, attitude, stress, and strain because of repeated tasks. Hence, it needs modern technological support to overcome the above pitfalls.

We can overcome the mentioned constraints using the embedded Artificial Intelligence driven camera to monitor the workers' tasks regularly and revert with feedback to do the right job on the first attempt itself.

II. METHODS AND TECHNIQUES

➢ *Dataset Creation, Including Pre-Processing & Augmentation:*

We have extracted each second frame from sample videos using Python. We also used the Roboflow (ak.3) algorithm for annotating each task of the operator and, created 2081 images with labels for further pre-processing and augmentation tasks. In the pre-processing stage, we applied auto-orientation, auto-resizing the images and automatically setting the contrast. In the augmentation process, we applied

- *Flip: Horizontal,90°*
- *Rotate: Clockwise, Counter-Clockwise,*
- *Rotation: Between -15° and +15°,*
- *Shear: ±15° Horizontal, ±15° Vertical,*
- *Gray-scale: Apply to 25% of images,*
- *Hue: Between -25° and +25°,*
- *Saturation: Between -25% and +25%,*
- *Brightness: Between -25% and +25%,*
- *Exposure: Between -25% and +25%,*
- *Blur: Up to 2.5px,*
- *Noise: Up to 5% of pixels,*
- *Mosaic: Applied,*
- *Bounding Box:*
- *Flip: Horizontal,*
- *Bounding Box: 90° Rotate: Clockwise, Counter-Clockwise,*
- *Bounding Box: Rotation: Between -15° and +15°,*
- *Bounding Box: Shear: ±15° Horizontal, ±15° Vertical,*
- *Bounding Box: Brightness: Between -25% and +25%,*
- *Bounding Box: Exposure: Between -25% and +25%,*
- *Bounding Box: Blur: Up to 2.5px,*
- *Bounding Box: Noise: Up to 5% of pixels.*

We verified every extracted sample video frame for considering the result of Improved Model Generalization, Accurate Representation, and Reduced Bias influencing factors.

ISSN No:-2456-2165

Augmentation techniques provide additional images for training, and this helps us in the object detection process, which further smoothens any environmental conditions.

In order to harness the capabilities of object detection techniques such as YOLO and address its limitations, we propose the integration of a novel Machine Learning workflow (ak.2) [Fig.2] that combines YOLO's efficiency with complementary methodologies. This workflow introduces a multi-stage approach, where YOLO's initial detections are further refined using advanced post-processing algorithms. In the following sections, we detail the components of this workflow and present experimental results that demonstrate its effectiveness in pushing the boundaries of object detection performance.

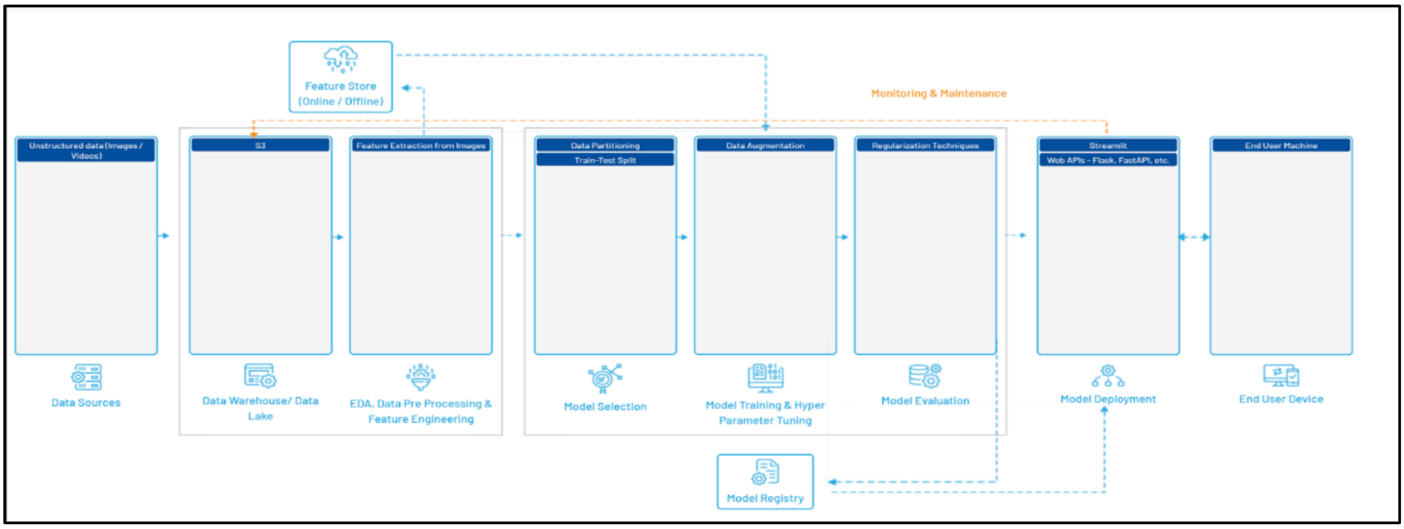


Fig 2 ML Workflow Architecture: A comprehensive overview of the machine learning pipeline for Clamping Sequence Detection. (Source: - Open-Source ML Workflow Tool- 360DigiTMG)

➢ *Model Architecture Diagram:*

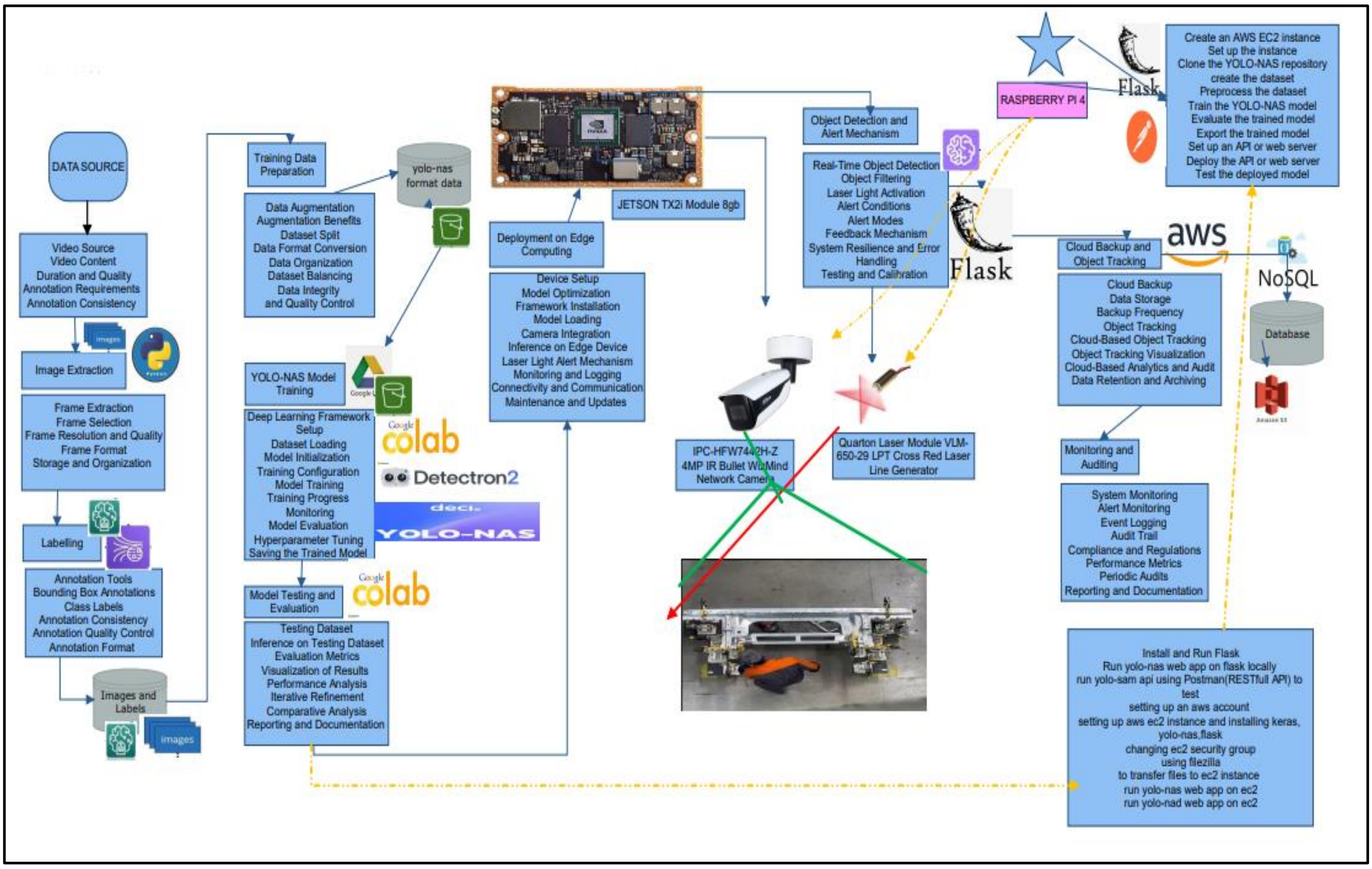


Fig 3 Architecture Diagram: Illustrating Data Flow and Components for Clamp Detection and Sequencing via Computer Vision and Machine Learning Methods.

ISSN No:-2456-2165

➢ *Role of Artificial Intelligence in Implementing the Poka-Yoke Principle.*

Japanese industrial engineers introduced Poka-Yoke [20] or mistake-proofing or error-proofing principles to create high-precision object or clamping status sequence detecting models.

There are two types of Poke-Yoke systems:- Warning Poka-Yoke [20] and Control Poka-Yoke [20] , which were implemented to ensure zero errors while doing clamping works. Artificial Intelligence (AI) plays a crucial role in preventing human errors and defects in manufacturing processes.

After several AI model comparisons, YOLO V8 [8,9] and YOLO-NAS [Fig.4] were the two models in this specific research.

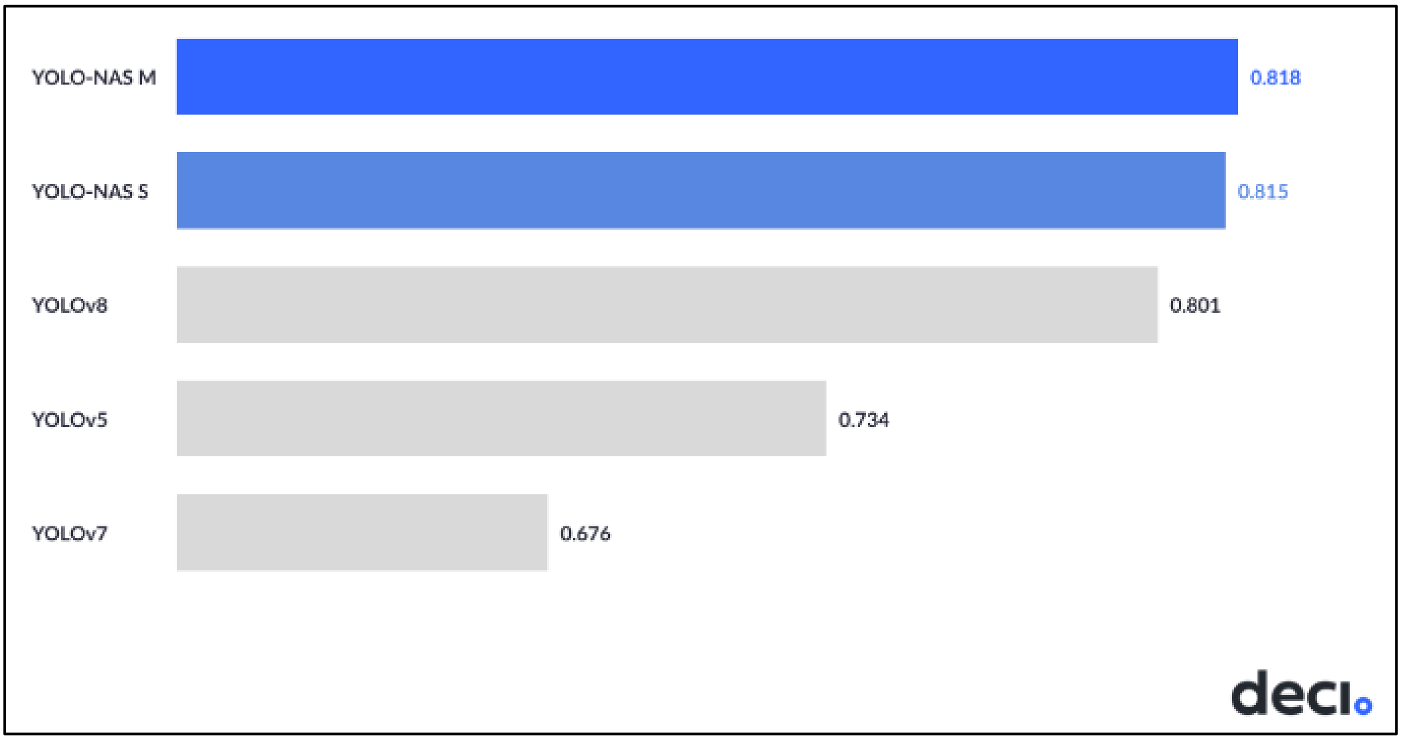


Fig 4 Yolo Models Comparison based on Average mAP Performance Sourced from the blog published by Augmented Startups.

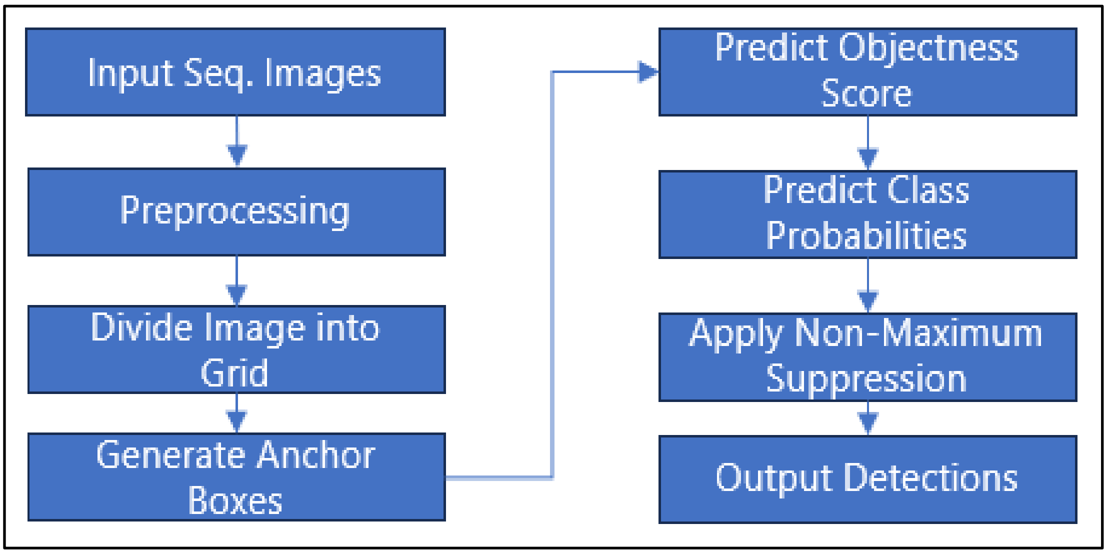


Fig 5 Object Detection Flowchart in Yolo model [1,2,3,4,5,6,10,12,13,14,15,16,17,18,19]

The aim is to detect the clamp's status, whether open or closed. 12 clamps open and 12 closed a total of 24 clamps. We trained our model on the dataset of 24 object detection in the first part.

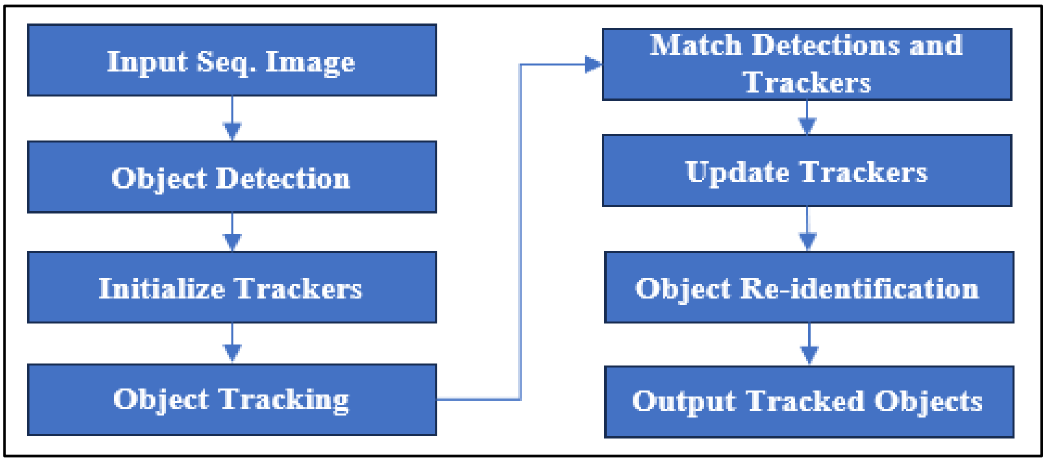


Fig 6 Object Tracking Flow chart in the YOLO model. [1,2,3,4,5,6,10,12,13,14,15,16,17,18,19]

The objects are tracked by first checking the status of the left-side clamps status and the status of the mirror image clamps on the right side, and then comparing both to track them for sequence pair analysis.

- *Clamping Pair Sequence*

Object detection and tracking are performed by an AI model, the operators are immediately alerted whether the sequence pair is right or wrong and the operators followed the right sequence every time.

- *API Interface*

The details of object detection and pair sequences are converted into JSON format. The JSON data is then updated on the AWS server for future audit and Total Quality Management (TQM) processes. Storing the data in JSON format on the AWS server enables efficient retrieval and analysis for auditing and TQM purposes.

- *Alert Mechanism*

Data received from the object detection and clamping pair sequence data will trigger in the API interface and will automatically alert the operator about the clamping task right or wrong sequence [Fig.3] through laser light over the assembly table, Raspberry Pi IV is used and connected with the laser diodes for alerting operators

III. RESULTS AND DISCUSSION

Based on the YOLO model's ability to track objects and infer spatial relationships can significantly enhance workflow monitoring and optimization [Fig.4] [Fig.5]. By analyzing manual operators clamping tasks captured in video feeds, YOLO can identify inefficiencies due to skipping sequences and deviations from standard operating procedures [Fig.6] [Fig.7] [Fig.8] [Fig.9] [Fig.10] [Fig.11]. This information enables us to streamline processes, by eliminating the re-doing the same tasks, boost the production speed, allocate the right resources effectively, and identify areas for improvement, leading to enhanced productivity and operational efficiency.

We have extracted video frames each second into 2081 images [Fig.12] [Fig.13] [Fig.14] [Fig.15] [Fig.16] [Fig.17]. These images we used to identify the tasks by marking labels for them with minute details also considered [Fig.18] [Fig.19].

Annotated labels we used for the input dataset for the YOLO model [Fig.20] [Fig.21]. We started working on an image dataset into all possibilities of camera capturing bottlenecks by augmentation process.

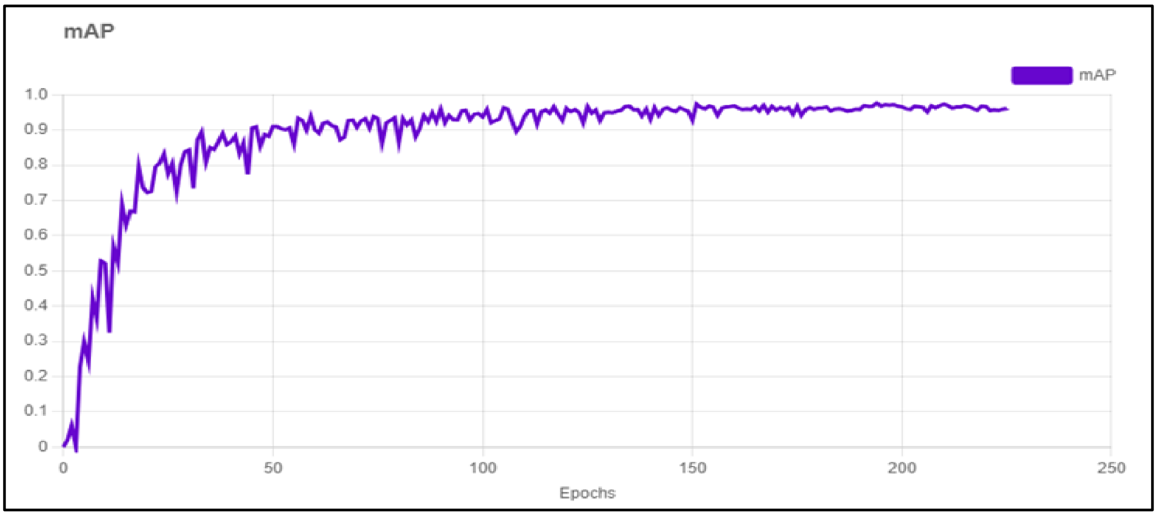


Fig 7 Yolo Model Mean Average Precision (mAP) Performance Data based on our Dataset.

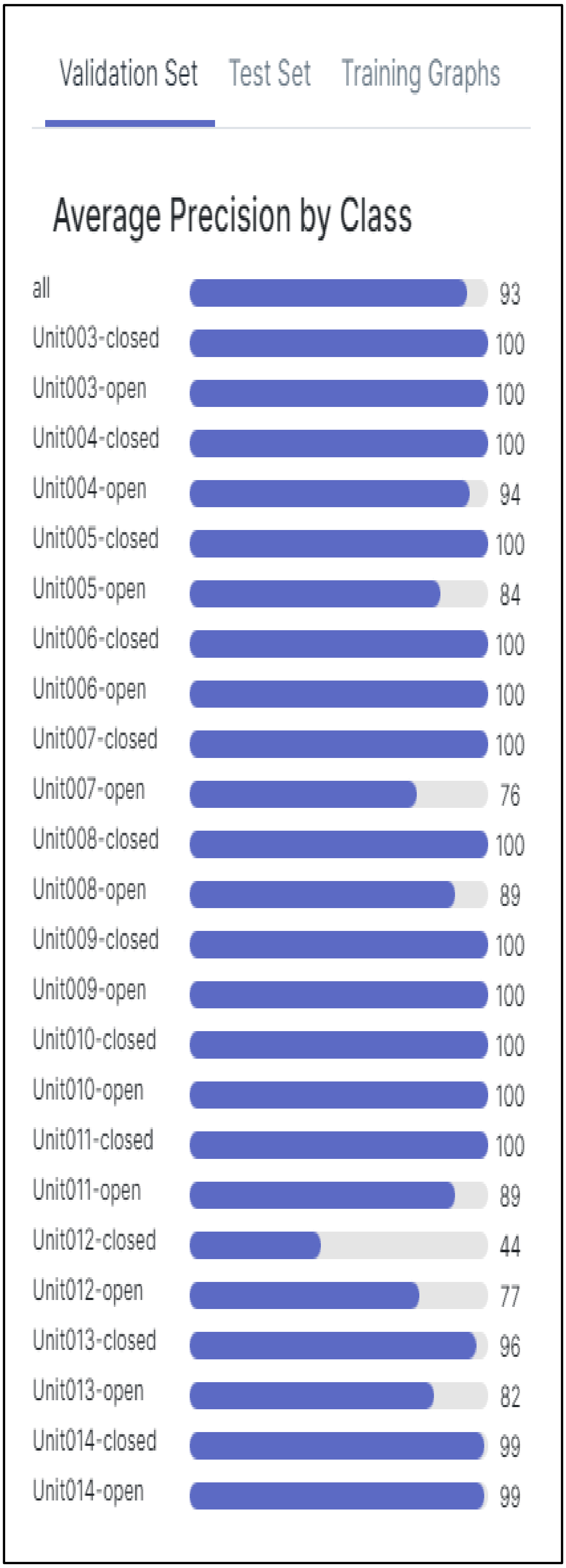


Fig 8 Validation Set Results for Average Precision, Compactly Conveying the Model's Performance.

Fig 9 Test Set Results for Average Precision are Depicted, Showcasing the Model's Performance.

ISSN No:-2456-2165

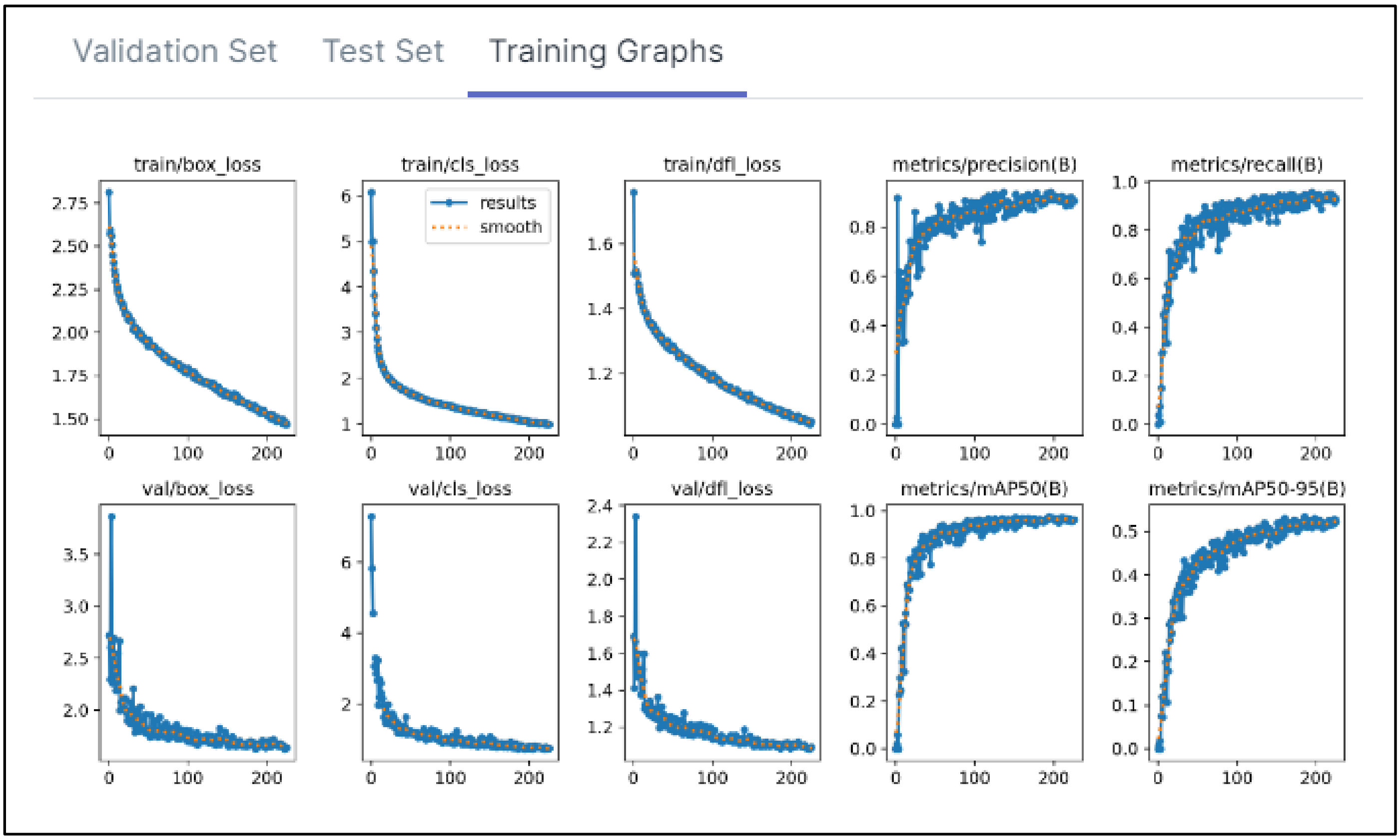


Fig 10 Training Graphs for the YOLO Model, Presenting its Learning Progress and Performance.

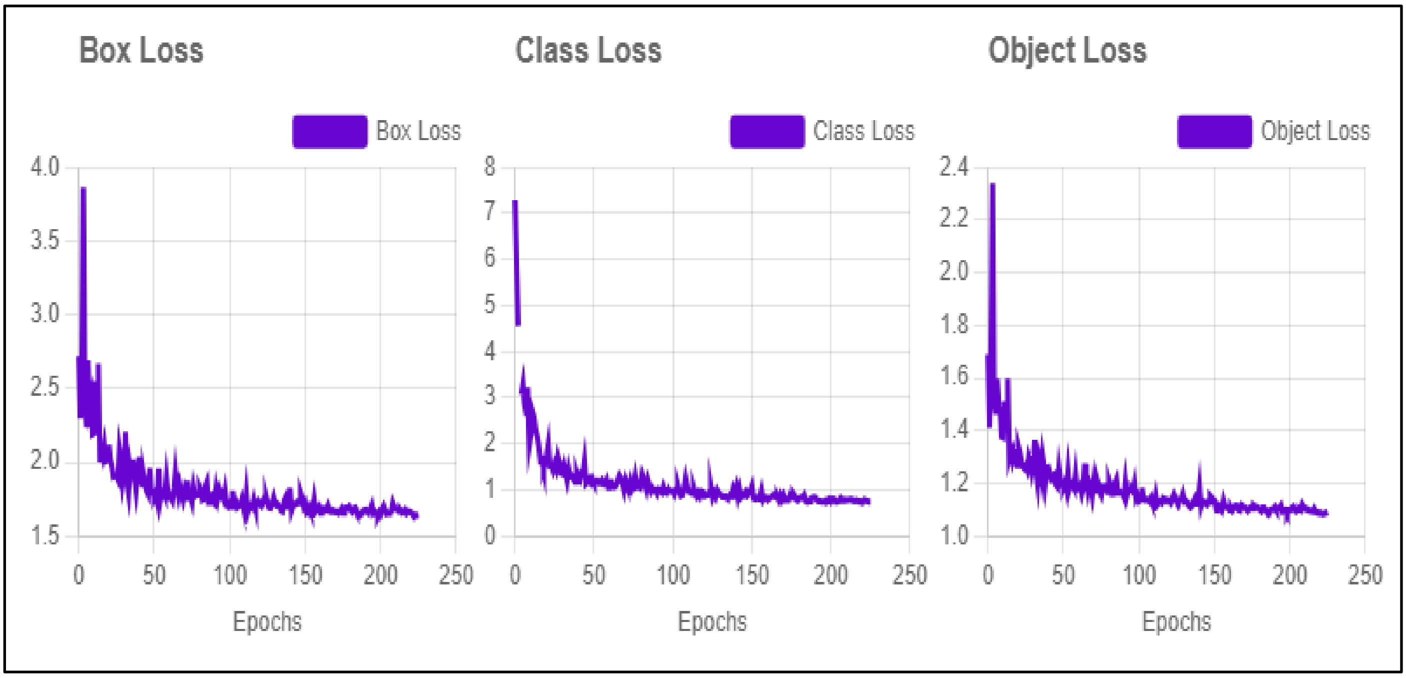
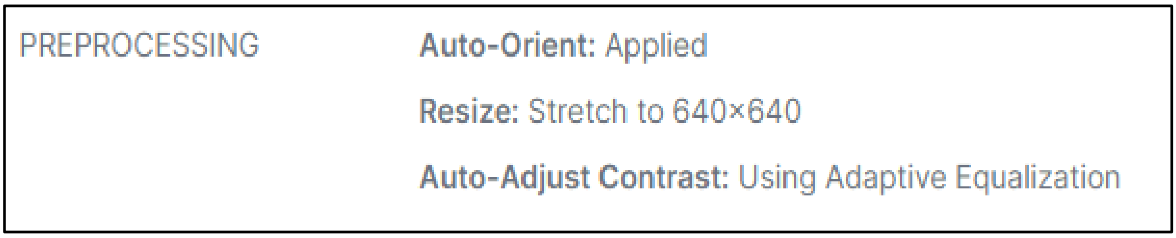


Fig 11 The Reduction of Object Detection Precision Losses via Improved Labeling of Boxes, Classes, and Objects



PREPROCESSING	**Auto-Orient:** Applied
	Resize: Stretch to 640×640
	Auto-Adjust Contrast: Using Adaptive Equalization

Fig 12 Visually Outlines Dataset Preprocessing, Highlighting Steps taken to Prepare the Data before it's used for Model Training

AUGMENTATIONS	**Outputs per training example:** 3
	Flip: Horizontal
	90° Rotate: Clockwise, Counter-Clockwise
	Rotation: Between -15° and +15°
	Shear: ±15° Horizontal, ±15° Vertical
	Grayscale: Apply to 25% of images
	Hue: Between -25° and +25°
	Saturation: Between -25% and +25%
	Brightness: Between -25% and +25%
	Exposure: Between -25% and +25%
	Blur: Up to 2.5px
	Noise: Up to 5% of pixels
	Mosaic: Applied
	Bounding Box: Flip: Horizontal
	Bounding Box: 90° Rotate: Clockwise, Counter-Clockwise
	Bounding Box: Rotation: Between -15° and +15°
	Bounding Box: Shear: ±15° Horizontal, ±15° Vertical
	Bounding Box: Brightness: Between -25% and +25%
	Bounding Box: Exposure: Between -25% and +25%
	Bounding Box: Blur: Up to 2.5px
	Bounding Box: Noise: Up to 5% of pixels

Fig 13 The Specifics of Image Augmentation Techniques applied to the Dataset for Enhanced Training Diversity

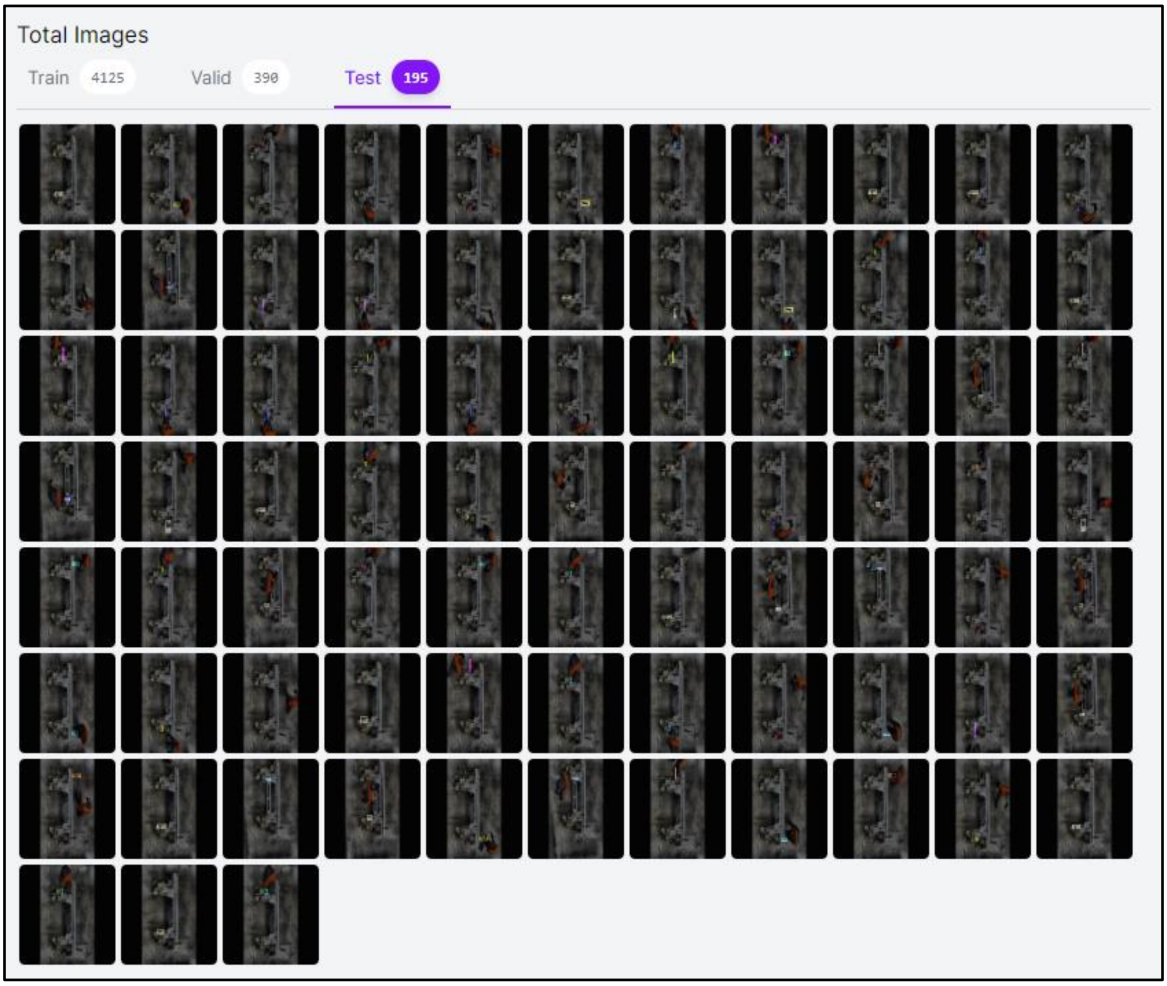


Fig 14 Train, Valid, and Test Views of the Image Dataset, Delineating Data Distribution Across Different sets.

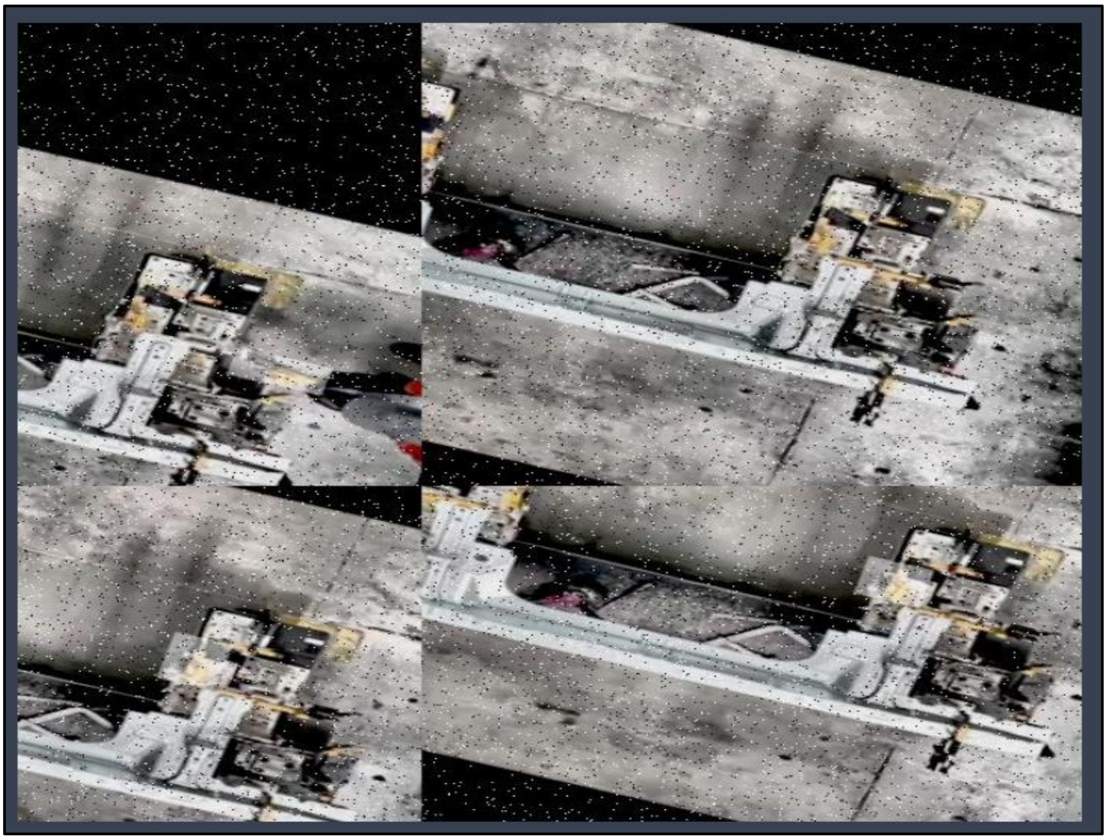


Fig 15 Image Augmentation Variations, Including Skew, Flip, and Contrast Transformations, Enhancing Dataset Diversity for Training.

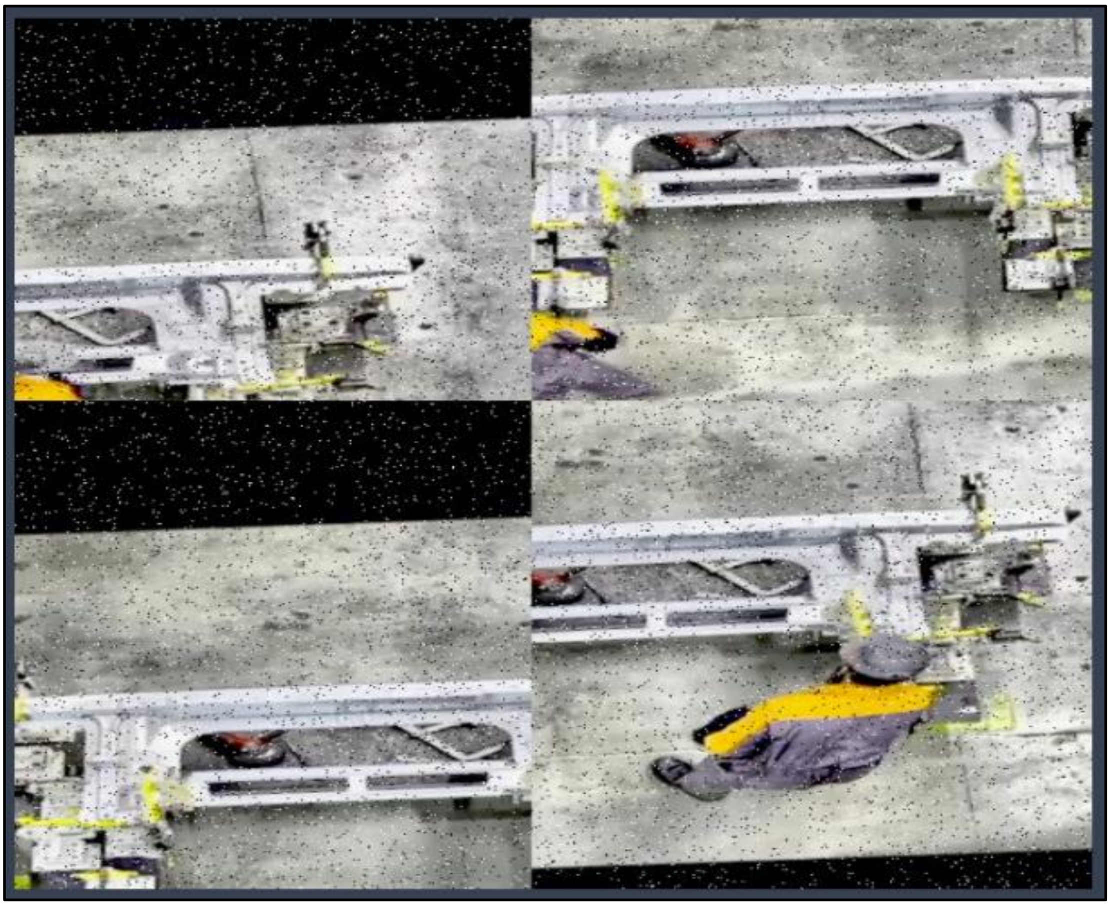


Fig 16 Image Augmentation Effects: Skew, Flip, and Contrast Variations, Enhancing Dataset Diversity and Training Robustness.

ISSN No:-2456-2165

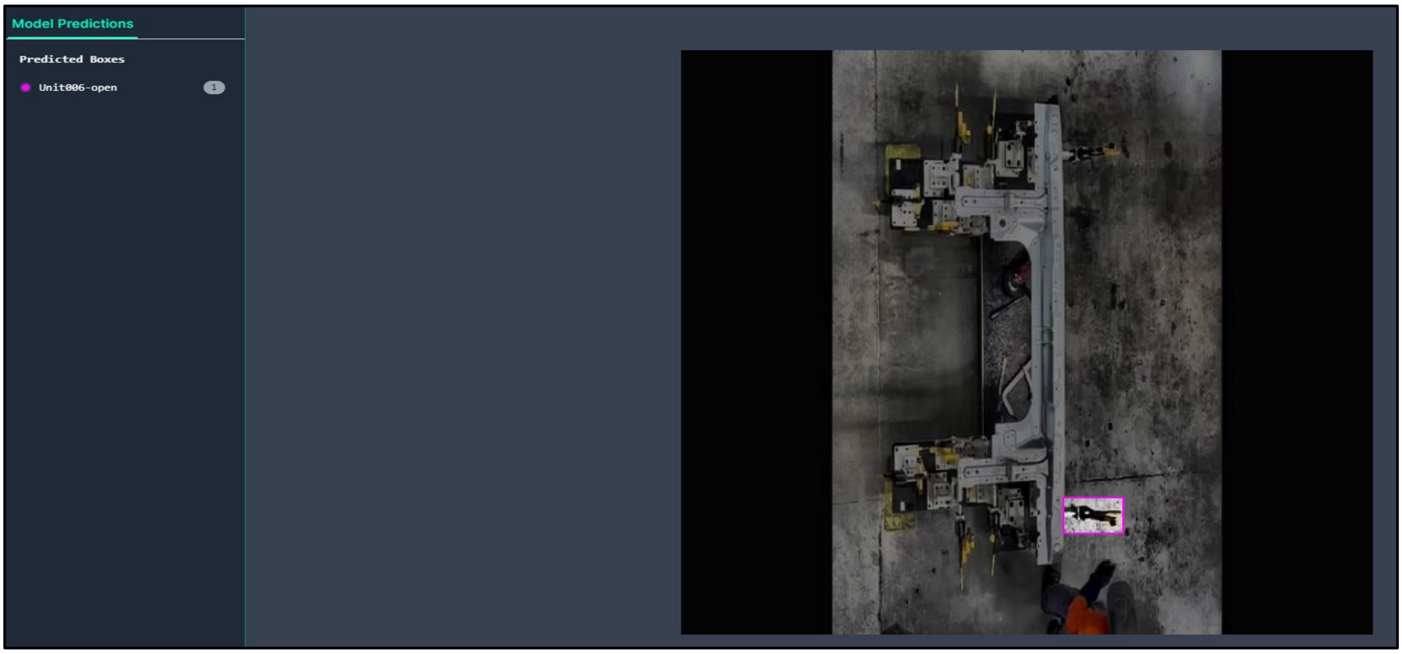


Fig 17 Views of Object Detection Bounding Boxes, Contributing to a Comprehensive Understanding of Detection Accuracy.

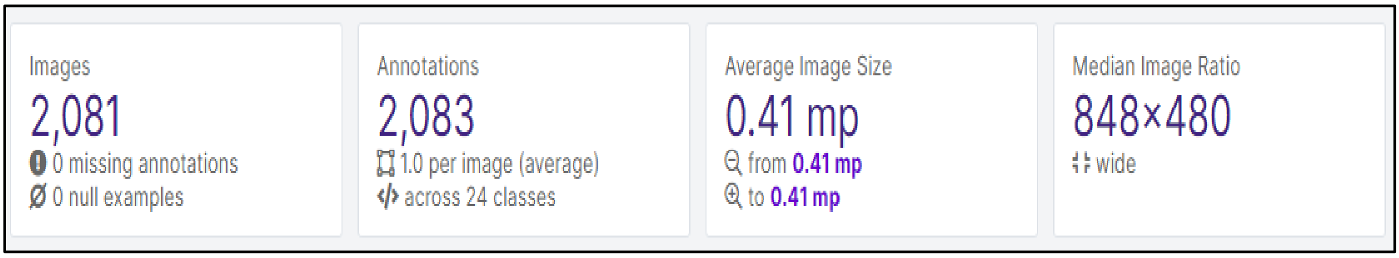


Fig 18 Dataset's Images, Annotations, and Average Image Sizes, Offering Insight into its Composition and Characteristics

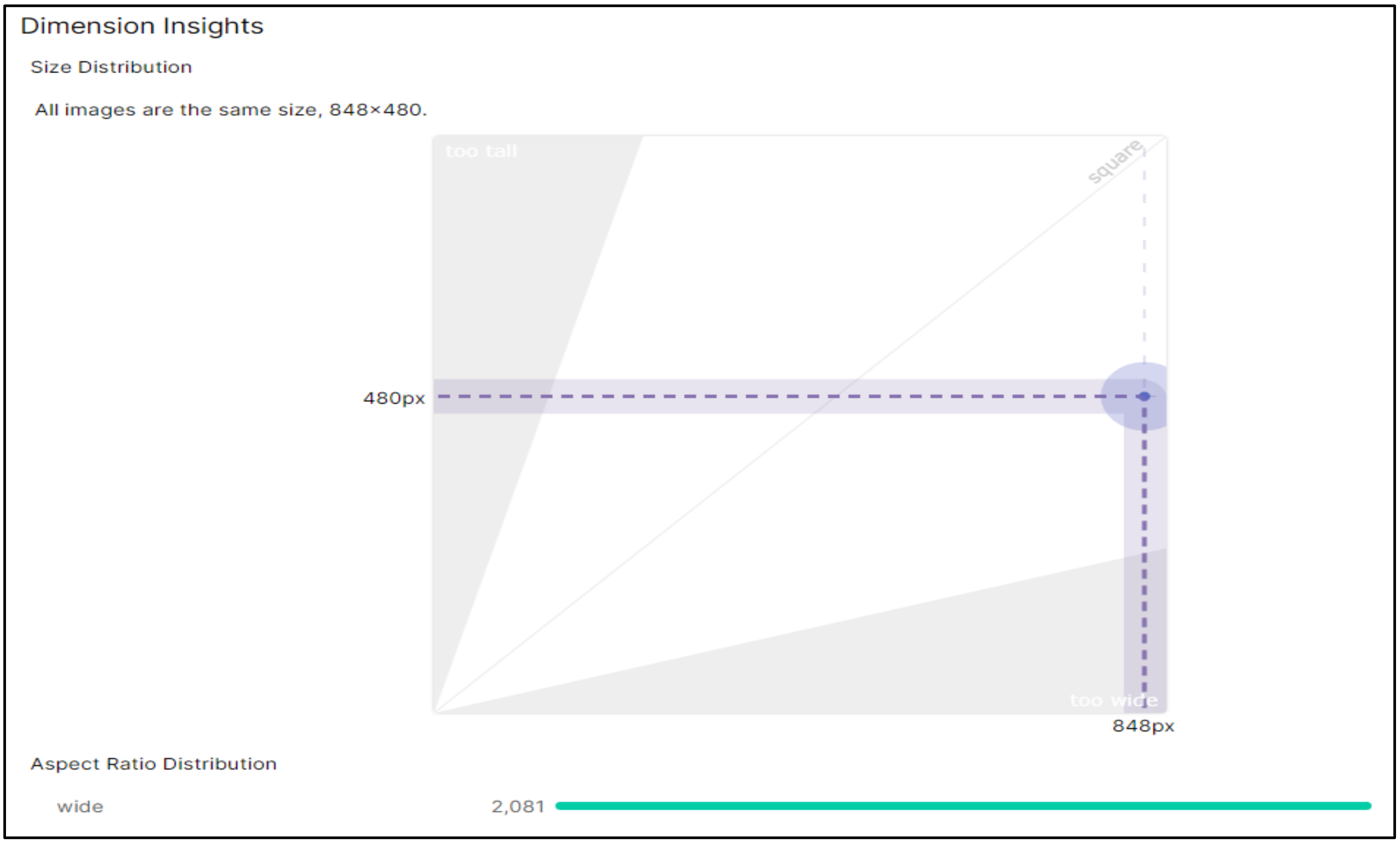


Fig 19 Distribution of Image Sizes within our Dataset, Highlighting Variations in Dimensions Across the Data.

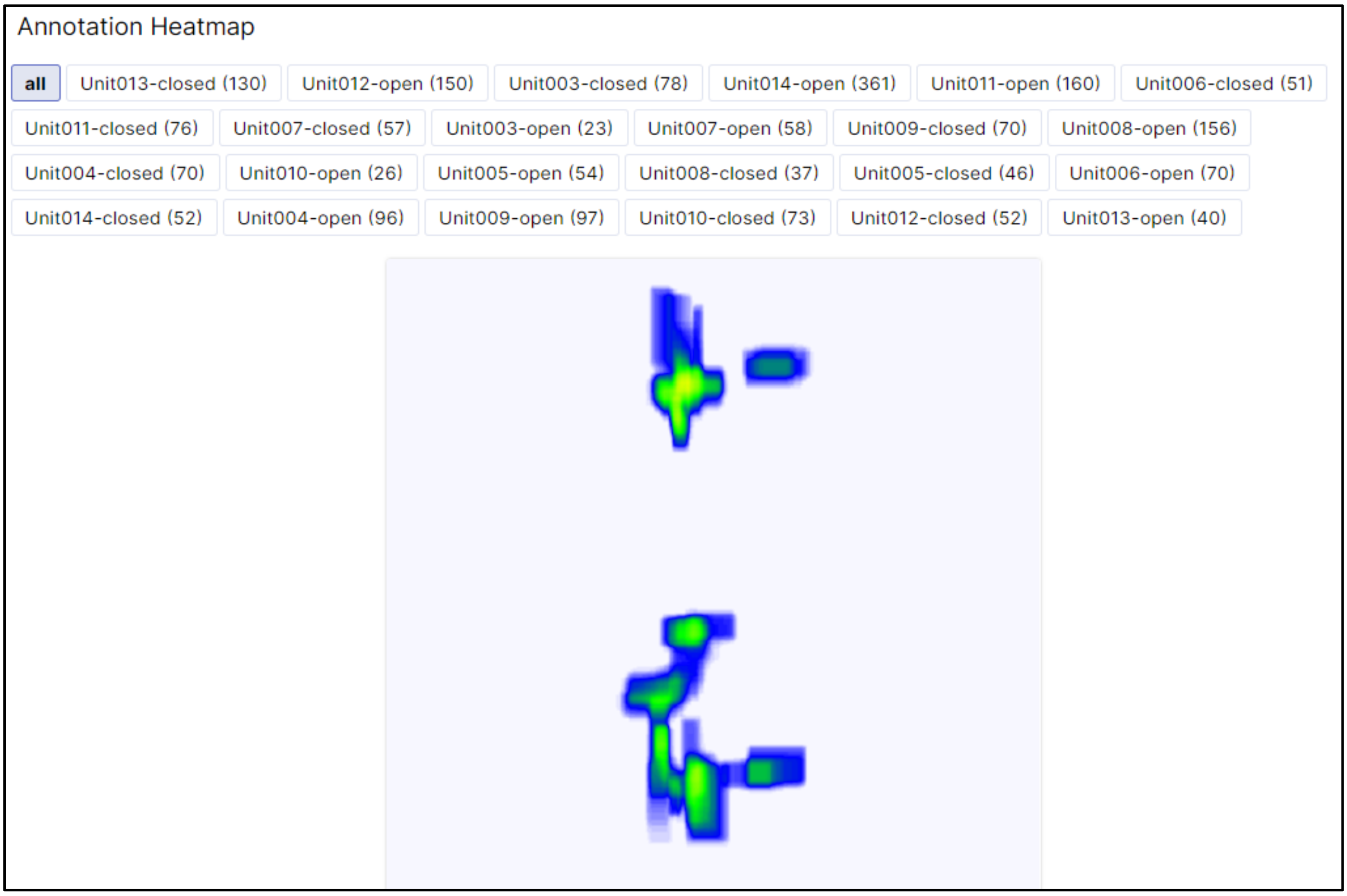


Fig 20 Annotation Heatmap, Providing a Visual Representation of the Density of Object Annotations within Images

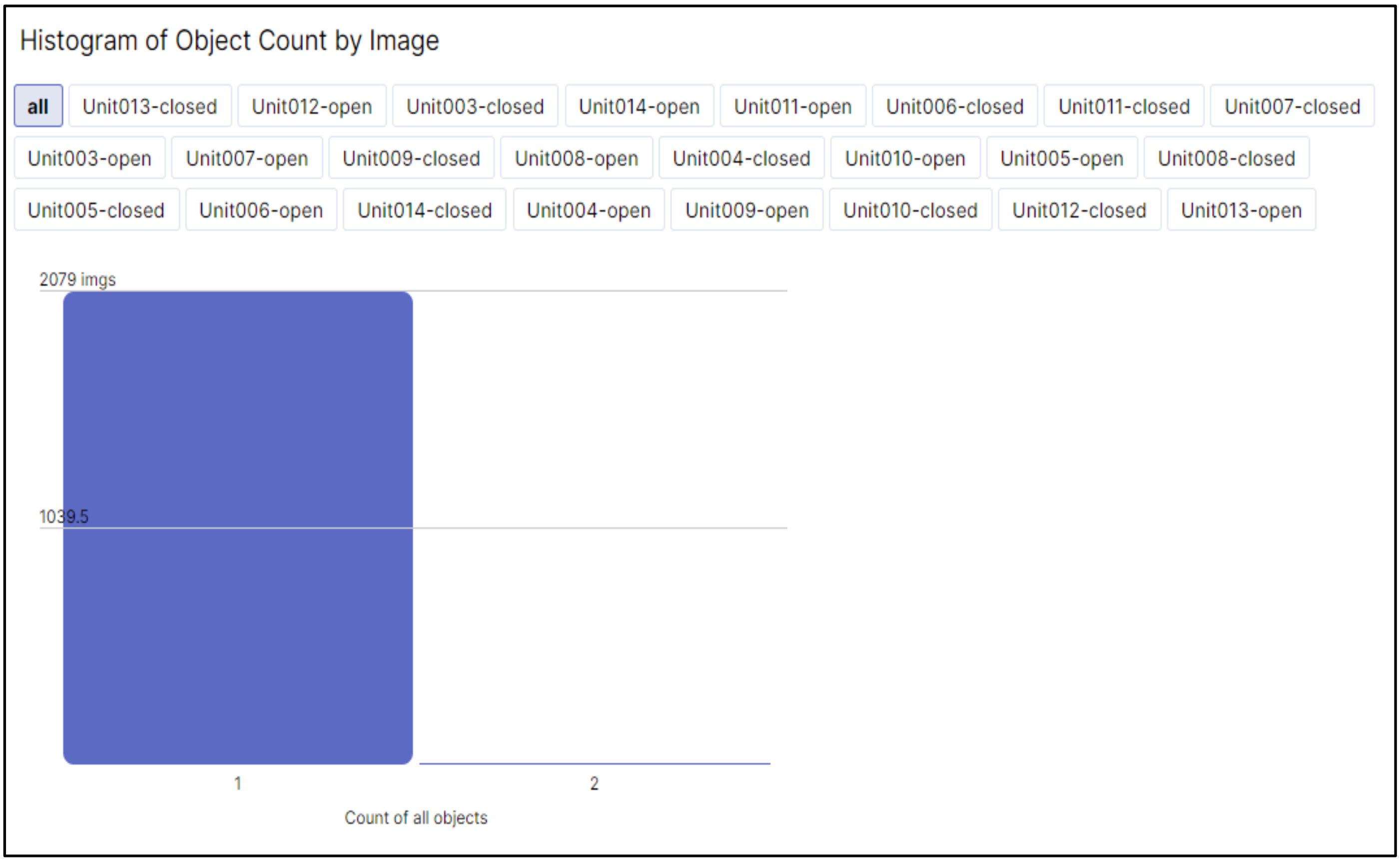


Fig 21 Histogram Illustrating how the Number of Objects per Image is Distributed Across the Dataset

This resulting model which we have developed for this research study can also be applied to different other manufacturing processes where the sequence of operations is crucial and requires human involvement such as Aerospace manufacturing, Tire assembly, Automotive body assembly line, Suspension system assembly and Printed circuit board (PCB) manufacturing where the risk of human errors can be reduced, ensuring consistent and accurate execution of tasks at each stage of the process.

IV. CONCLUSION

Modern AI plays a crucial role in the automobile manufacturing manual clamping process, monitoring and timely alerting the human errors so that we achieve the nearest zero human error in the truck body parts assembling stage. This will speed up the manufacturing process, nullify repetitive tasks, and boost production efficiency.

- *Declarations*

V. ACKNOWLEDGMENTS

- We acknowledge that with the consent from 360DigiTMG, we have used the CRISP-ML(Q) methodology (ak.1) and the ML Workflow which are available as open-source in the official website of 360DigiTMG (ak.2).
- We acknowledge that for the Data Pre-processing, we have used the Roboflow tool (free version), the link to the tool is Roboflow application for Data Pre-processing, https://docs.roboflow.com/image-transformations/image-augmentation.(ak.3).

➢ *Funding and Financial Declarations:*

- *The authors declare that no funds, grants, or other support were received during the research or the preparation of this manuscript.*
- *The authors declare that they have no relevant financial or non-financial interests to disclose.*

➢ *Author Contributions:*

Author Bharani Kumar Depuru has conceptualized the research and development of the model and provided with the necessary technical resources, also guided and directed the study. Bhushan Sharad Patil led the research and contributed to the analysis and model development. Material preparation, data collection and analysis were performed by Amrathakara Bhat and Aishwarya Dhadd under the guidance and mentorship of Bhushan Sharad Patil and Bharani Kumar Depuru. The first draft of the manuscript was written by Amrathakara Bhat and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

➢ *Data Availability Statement:*

The datasets used, generated and/or analysed during this study are not publicly available due to internal Data Privacy Policy but are available from the corresponding author on reasonable request.

➢ *Compliance with Ethical Standards*

- *Disclosure of Potential Conflicts of Interest:*

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

- *Research Involving Human Participants and/or Animals:*

It is declared by all the authors that there was no involvement of any human and/or animal trial or test in this research.

- *Informed Consent:*

As there were no human subject involved in this research hence a informed consent is not applicable to the best of the authors' understanding.

- *Conflict of Interest Statement:*

The authors declare that there are no conflicts of interest that could influence the results or interpretation of this manuscript. This research was conducted in an impartial and unbiased manner, and there are no financial, personal, or professional relationships that might be perceived as having influenced the content or conclusions presented in this work.

REFERENCES

[1]. Ercihan Kiraci, Arnab Palit, Alex Attridge & Mark A. Williams (2022) The effect of clamping sequence on dimensional variability of a manufactured automotive sheet metal sub-assembly, International Journal of Production Research, DOI: 10.1080/ 00207543.2022.2152897

[2]. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788). https://doi.org/10.48550/arXiv.1506.02640.

[3]. Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271). https://doi.org/ 10.48550/arXiv.1612.08242.
Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767. https://doi.org/10.48550/arXiv. 1804.02767.

[4]. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934. https://doi.org/10.48550/arXiv.2004.10934.

[5]. Benjumea, A., Teeti, I., Cuzzolin, F., & Bradley, A. (2021). YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles. arXiv preprint arXiv:2112.11798. https://doi.org/10.48550/arXiv. 2112.11798.

[6]. A hybrid pixel-based background model for image foreground object detection in complex sence, by Chung-chi Lin; Wen-kai Tsai; Ming-hwa Sheu,INSPEC Accession Number: 12908747,Date Added to IEEE Xplore: 02 August 2012,Date of Conference: 03-04 July 2012, Prague, Czech Republic, https://doi.org/10.1109/TSP.2012.6256391.

[7]. High-accuracy background model for real-time video foreground object detection, Optical Engineering 51(2), 027202 (February 2012), by Wen-Kai Tsai, Chung-Chi Lin, Ming-Hwa Sheu, Paper 110960 received Aug. 9, 2011; revised manuscript received Nov. 16, 2011; accepted for publication Dec. 12, 2011; published online Mar. 2, 2012. https://doi.org/10.1117/1.OE.51.2.027202

[8]. Tracking nautical objects in real-time via layered saliency detection Paper by Matthew Dawkins, Zhaohui Sun, Arslan Basharat, Amitha Perera, Anthony Hoogs Proceedings Volume Geospatial InfoFusion and Video Analytics IV; and Motion Imagery for ISR and Situational Awareness II, 908903 (2014) https://doi.org/10.1117/12.2058122.

[9]. Robust background modeling for enhancing object tracking in video Paper by Richard J. Wood, David Reed, Janet Lepanto, John M. Irvine, Proceedings Volume Geospatial InfoFusion and Video Analytics IV; and Motion Imagery for ISR and Situational Awareness II, 908902 (2014). https://doi.org/10.1117/12.2047258.

[10]. Feature fusion and label propagation for textured object video segmentation Paper V. B. Surya Prasath, Rengarajan Pelapur, Kannappan Palaniappan, Gunasekaran Seetharaman, Proceedings Volume Geospatial InfoFusion, and Video Analytics IV; and Motion Imagery for ISR and Situational Awareness II, 908904 (2014) https://doi.org/10.1117/12.2052983.

[11]. The effect of state dependent probability of detection in multitarget tracking applications Paper by A. Gning, W. T. L. Teacy, R. V. Pelapur, H. AliAkbarpour, K. Palaniappan, G. Seetharaman, S. J. Julier Proceedings Volume Geospatial InfoFusion and Video Analytics IV; and Motion Imagery for ISR and Situational Awareness II, 908905 (2014) https://doi.org/10.1117/12.2053968.

[12]. Vehicle change detection from aerial imagery using detection response maps Paper by Zhaohui H. Sun, Mathew Leotta, Anthony Hoogs, Rusty Blue, Robert Neuroth, Juan Vasquez, Amitha Perera, Matthew Turek, Erik Blasch Proceedings Volume Geospatial InfoFusion and Video Analytics IV; and Motion Imagery for ISR and Situational Awareness II, 908906 (2014) https://doi.org/10.1117/12.2055362.

[13]. Weakly Supervised Video Salient Object Detection, via Point Supervision, by Shuyong Gao, Haozhe Xing, Wei Zhang, Yan Wang, Qianyu Guo, Wenqiang Zhang, Shanghai Key Laboratory of Intelligent Information Processing, School of Computer Science, Fudan University, Shanghai, China. https://doi.org/10.48550/arXiv.2207.07269.

[14]. Pyramid Dilated Deeper ConvLSTM for Video Salient Object Detection by Hongmei Song, Wenguan Wang1 , Sanyuan Zhao, Jianbing Shen,and Kin-Man Lam,Springer Nature Switzerland AG 2018,V. Ferrari et al. (Eds.): ECCV 2018, LNCS 11215, pp. 744–760, 6th October 2018. https://doi.org/10.1007/978-3-030-01252-6_44.

[15]. Online Training of Discriminative Parameter for Object Tracking-by-Detection in a Video, by Vijay Kumar Sharma, Bibhudendra Acharya, and K. K. Mahapatra, Springer Nature Singapore Pte Ltd. 2019, J. Nayak et al. (eds.), Soft Computing in Data Analytics, Advances in Intelligent Systems and Computing 758, Apil 2020 Nanoelectronics, Circuits and Communication Systems (pp.355-369). http://dx.doi.org/10.1007/978-981-15-2854-5_31.

[16]. Video Analytics Algorithm For Detecting Objects Crossing Lines In Specific Direction Using Blob-Based Analysis, by Zulaikha Kadim , Liang Kim Men g, Norshuhada Samud in, Khairunnisa Mohamed Joari , Khairil Hafriza, Choong Teck Lion g, Hon Hock Woon.MoMM '12: Proceedings of the 10th International Conference on Advances in Mobile Computing & MultimediaDecember 2012Pages 306–312, 3rd December 2012. https://doi.org/10.1145/2428955.2429016.

[17]. Pyramid Constrained Self-Attention Network for Fast Video Salient Object Detection Yuchao Gu, Lijuan Wa ng, Ziqin Wang, Yun Liu, Ming-Ming Cheng, Shao-Ping Lu, April 2020Proceedings of the AAAI Conference on Artificial Intelligence 34(07):10869-10876, http://dx.doi.org/10.1609/aaai.v34i07.6718.

[18]. Attention Embedded Spatio-Temporal Network for Video Salient Object Detection, Lili Huang, Pengxiang Yan, Guanbin Li, Qing Wang, Liang Lin, Page(s): 166203 - 16621 3, Date of Publication: 12 November 219, Electronic ISSN: 2169-3536, INSPEC Accession Number: 19144213. https://doi.org/10.1109/ACCESS.2019.2953046.

[19]. A Simple Objective Method for Automatic Error Detection in Stereoscopic 3D Video, Shirish Sharma, Eva Cheng, Ian Burnett, Date of Conference: 22-25 September 2015, Date Added to IEEE Xplore: 02 November 2015, INSPEC Accession Number: 15573035. https://doi.org/10.1109/BDVA.2015.7314285.

[20]. Romero, D., Gaiardelli, P., Powell, D.J., Zanchi, M. (2022). Intelligent Poka-Yokes: Error-Proofing and Continuous Improvement in the Digital Lean Manufacturing World. In: Kim, D.Y., von Cieminski, G., Romero, D. (eds) Advances in Production Management Systems. Smart Manufacturing and Logistics Systems: Turning Ideas into Action. APMS 2022. IFIP Advances in Information and Communication Technology, vol 664. Springer, Cham.