SEBHA UNIVERSITY CONFERENCE PROCEEDINGS, VOL.03 NO. 2 2024
DOI: 10.51984/SUCP.V3I2.3409

وقائع مؤتمرات جامعة سبها
Sebha University Conference Proceedings

Confrence Proceeeding homepage: http://www.sebhau.edu.ly/journal/CAS

Improving Vehicle Identification Number Detection Accuracy with YOLOv5 and Histogram Equalization

* Hasan Alkhadafe [1], Zahiyah Khalleefah[2], Ibrahim Nasir[3]

[1]Department of Information Systems Faculty of Information Technology, Sebha University

[2]Department of Artificial Intelligence, Sebha University

[3] Department of Electrical &Electronic Engineering, Sebha University

Keywords:
Grayscale
Histogram Equalization
Image Enhancement
VIN
YOLOv5

ABSTRACT

This study examines the effectiveness of different image preprocessing techniques for object detection models, using a dataset of VIN images from Roboflow. The dataset was segmented into training, validation, and testing subsets, encompassing a range of conditions such as noise, rain, varying lighting, and reflections. Model performance was evaluated through metrics including precision, recall, average precision (AP), mean average precision (mAP), error rate reduction, and frames per second (FPS).The baseline model, trained on the original dataset, achieved a precision of 97.9% and a recall of 95.7%, with an mAP@0.5 of 99.1% but a lower mAP@0.5:0.95 of 62.3%. Applying Histogram Equalization (HE) resulted in improved recall but reduced precision, with mAP@0.5:0.95 values remaining comparable to the original dataset. The HE+RGB preprocessing showed minor performance changes, with inconsistent improvements in recall and precision. Adaptive Histogram Equalization (AHE) notably improved model performance, reaching a precision of 98.8% and recall of 99.6%, with mAP@0.5 and mAP@0.5:0.95 values of 74.3%, 77.0% respectively. The CLAHE preprocessing technique outperformed all others, achieving the highest precision (99.4%), recall (98.6%), and mAP@0.5:0.95 (75.2% in training, 77.9% in validation, and 75.2% in testing), demonstrating the best balance of accuracy and generalization with minimal misclassifications. Overall, CLAHE emerged as the most effective preprocessing method, offering superior performance across all evaluation metrics.

تحسين دقة اكتشاف رقم تعريف المركبة باستخدام YOLOv5 ومعادلة الرسم البياني .

*حسن صالح القذافي[1] و زاهية شحات جعراني خليفة[2] و إبراهيم السنوسي نصر[3]

[1]قسم نظم المعلومات كلية تكنولوجيا المعلومات، جامعة سبها

[2]قسم الذكاء الاصطناعي، جامعة سبها

[3] قسم الهندسة الكهربائية والالكترونية، جامعة سبها

الكلمات المفتاحية:
التدرج الرمادي
تحسين الصور
معادلة الرسم البياني
VIN
YOLOv5

الملخص

تستعرض هذه الدراسة فعالية تقنيات المعالجة المسبقة المختلفة لنماذج الكشف عن الأجسام، باستخدام مجموعة بيانات تحتوي على 1145 صورة لرقم تعريف المركبات (VIN) من Roboflow. تم تقسيم مجموعة البيانات إلى مجموعات تدريب، تحقق، واختبار، وشملت مجموعة من الظروف مثل الضوضاء، الأمطار، اختلافات الإضاءة، والانعكاسات. تم تقييم أداء النموذج من خلال مقاييس تشمل الدقة، الاسترجاع، الدقة المتوسطة(AP) ، الدقة المتوسطة العامة(mAP) ، تقليل معدل الأخطاء، والإطارات في الثانية (FPS).حقق النموذج الأساسي، الذي تم تدريبه على مجموعة البيانات الأصلية، دقة قدرها 97.9% واسترجاع قدره 95.7%، مع قيمة mAP@0.5 قدرها 99.1% ولكن قيمة mAP@0.5:0.95 كانت

*Corresponding author:
E-mail addresses: has.alkhadafe@sebhau.edu.ly ,(Z. Khalleefah) za.khalifa@sebhau.edu.ly ,(I. Nasir) ibr.nasir@sebhau.edu.ly
Article History : Received 20 June 2024 - Received in revised form 24 September 2024 - Accepted 06 October 2024

أقل عند 62.3%. أدى تطبيق تقنية التعديل الهيستوجرام (HE) إلى تحسين الاسترجاع ولكن مع تقليل الدقة، حيث بقيت قيم mAP@0.5:0.95 مشابهة بمجموعة البيانات الأصلية. أظهرت تقنية HE+RGB تغييرات طفيفة في الأداء، مع تحسينات غير متسقة في الاسترجاع والدقة، وقدرة عامة أقل .حقق التعديل الهيستوجرام التكيفي (AHE) تحسينًا ملحوظًا في أداء النموذج، حيث وصلت الدقة إلى 98.8% والاسترجاع إلى 99.6%، مع قيم mAP@0.5:0.95 بلغت 74.3% و77.0%. أما تقنية CLAHE فقد تفوقت على جميع التقنيات الأخرى، محققة أعلى دقة (99.4%) واسترجاع (98.6%)، وقيم mAP@0.5:0.95 (75.2% في التدريب، 77.9% في التحقق، و75.2% في الاختبار)، مما يظهر أفضل توازن بين الدقة والقدرة على التعميم مع الحد الأدنى من الأخطاء. بوجه عام، برزت تقنية CLAHE كأكثر تقنيات المعالجة المسبقة فعالية، حيث قدمت أداءً ممتازًا عبر جميع مقاييس التقييم.

1. Introduction

Recent advancements in object detection models have revolutionized the field of computer vision, by enabling the development of sophisticated computer vision models capable of complex tasks such as object detection with high accuracy. One area where this technology has shown great potential is in the automotive industry, particularly in the identification of Vehicle for security, safety, and tracking purposes, A Vehicle Identification Number (VIN), stands as the unique and permanent alphanumeric code assigned to each motor vehicle during the manufacturing process. This 17-character identifier functions as the car's individual fingerprint, enabling the precise tracking and distinction of every unit produced. Notably, the VIN composition excludes the letters I, O, Q, and Z to eliminate any potential confusion that might arise between numerals and alphabetic characters[1]. YOLO (You Only Look Once) is a widely used object detection system in the fields of computer vision and image processing. YOLO is particularly known for its real-time capabilities, as it can detect objects in an image with both high accuracy and speed[2][3]. By leveraging YOLOv5, this model can accurately identify and extract VINs from images, providing a valuable tool for law enforcement, insurance companies, and other stakeholders in the automotive industry. However, one crucial aspect is the impact of pre-processing techniques on model performance. Histogram equalization is a widely used method to enhance the contrast of images, thereby improving object detection accuracy. This research aims to investigate the effects of applying histogram equalization to the Vehicle Identification Number (VIN) dataset before feeding it into YOLOv5. By comparing the performance metrics of YOLOv5 before and after the application of histogram equalization, we can gain insights into the potential benefits of this pre-processing technique. This study seeks to contribute valuable insights to the ongoing discourse on optimizing object detection models for improved accuracy and efficiency.

2. RELATED WORKS

- **IMAGE ENHANCEMENT**

The image enhancement process focuses on elevating an image's visual quality, making it more aesthetically pleasing, or facilitating the extraction of valuable information. To achieve these goals, a diverse arsenal of techniques are employed, broadly categorized into two distinct approaches: spatial and frequency domain[4]. Spatial domain techniques work directly on the individual building blocks of an image, and its pixels, manipulating their values to achieve the desired visual improvements. Frequency domain techniques delve deeper, transforming the image into a different representation based on its fundamental frequencies. By manipulating these frequencies, targeted enhancements can be made to specific aspects of the image. This approach utilizes techniques like Fourier transforms. Image enhancement offers a toolbox for addressing specific visual limitations, encompassing techniques like brightness and contrast adjustment, histogram equalization, noise reduction, sharpening, and color correction. These methods find application not just in improving aesthetics but also in various fields like photography, medical imaging, satellite imaging, and video processing[5].

- **Histogram equalization (HE)**

Histogram equalization (HE) emerges as a prominent technique for contrast enhancement. HE operates by manipulating the distribution of pixel intensities within an image. This distribution aims to achieve a more uniform spread of pixel values across the entire spectrum of possible intensity levels[6]. In HE. This distribution is first obtained by computing the image's histogram, a graphical representation that depicts the frequency of occurrence for each pixel value[7]. Following, HE proceeds by applying a redistributing to the pixel values. This redistributing aims to achieve a more uniform distribution of these values across the entire spectrum of possible intensities. In essence, HE redistributes the pixel intensities within the image to create a more balanced representation[8]. The HE comprises two fundamental operations. Firstly, the cumulative distribution function (CDF) of the image's histogram is computed. Subsequently, the pixel values of the input image are mapped to a novel range leveraging the derived CDF[9]. HE is a versatile image enhancement technique applicable to both grayscale and color images. The HE formula is: The histogram of a digital image with intensity levels in the range [0, L-1] n_k is the number of pixels in the image denoted by a matrix M×N, where M and N are the row and column dimension of the image.

$$P(k) = \frac{n(k)}{M \times N}, for\ k = 0,1 \dots L-1$$

Compute (CDF) of the normalized histogram with Formula.

$$CDF(j) = SUM\big(p(k)\big), for\ k = 0,1, \dots, j \quad and\ j = 0,1, \dots, L-1$$

Compute the new intensity values for each pixel.

$$s = round\ (L-1) \times CDF(I)$$

The histogram equalization (HE) process involves substituting each pixel's original intensity value in the input image with its corresponding new intensity, as determined by the HE formula. This formula maps the original intensity values to novel values based on the cumulative distribution function of the image's histogram, thereby dispersing the intensity values across the full range of available intensities. Consequently, the output image exhibits enhanced global contrast.

- **Adaptive Histogram Equalization (AHE)**

Adaptive Histogram Equalization (AHE) is an enhancement technique derived from traditional histogram equalization, specifically designed to improve contrast in areas with low contrast or uneven illumination. Unlike standard histogram equalization, which applies a global transformation to the entire image, AHE operates locally. It divides the image into smaller, non-overlapping tiles and applies histogram equalization to each tile individually. This localized approach allows AHE to enhance contrast more effectively across various regions with different brightness levels or contrast [10]. In AHE, each tile's histogram is computed and used to adjust the pixel intensities based on the cumulative distribution function (CDF) of that region's histogram. The tiles are then combined using interpolation, which ensures smooth transitions between adjacent regions [10]. This method significantly enhances local contrast, making it particularly useful for images where fine detail visibility is critical

- **Contrast Limited Adaptive Histogram Equalization (CLAHE)**

Contrast Limited Adaptive Histogram Equalization (CLAHE) was developed to address the problem of noise amplification in AHE by introducing a contrast-limiting mechanism. CLAHE limits the

amplification of contrast in homogeneous areas by clipping the histogram at a predefined threshold, redistributing the excess values across the intensity range [11]. This approach reduces the over-enhancement of noise and produces more visually appealing results, making it particularly effective in medical imaging and low-light conditions [12][13].

CLAHE operates similarly to AHE, but with an added step to limit contrast:

a. Divide the image into tiles as in AHE.
b. Clip the histogram at a predefined threshold to limit noise amplification.
c. Redistribute pixel intensities using the clipped histogram to compute the CDF for each tile.
d. Map the pixel intensities and merge the tiles smoothly using bilinear interpolation.

By capping the histogram, CLAHE ensures that the image's contrast enhancement does not become excessive, while still improving local contrast. CLAHE's ability to balance noise suppression with contrast enhancement has led to its widespread use in object detection models, such as YOLOv5, as it improves accuracy while mitigating the risk of false positives and negatives [14].

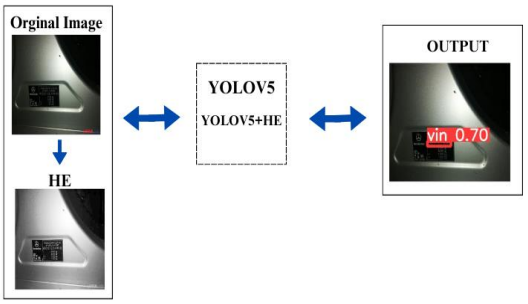
- **You Only Look Once (YOLO)**

The YOLO object detection algorithm has undergone a series of iterative developments, including version YOLOv1, YOLOv2, YOLOv3, and YOLOv4 [15]. These successive iterations have incorporated enhancements to the original algorithm, including increased accuracy, accelerated processing velocity, and enhanced capabilities for detecting small objects and managing occluded objects [16]. YOLOv5, was created by Ultralytics and launched by Glenn Jocher in 2020. Similar to earlier YOLO versions, YOLOv5 is built on the EfficientDet object detection framework and has a single-stage detector (SSD) architecture. YOLOv5's speed and accuracy are two major upgrades over earlier iterations. Because YOLOv5 uses a more effective backbone network and better training methods than its predecessors, it is faster and more accurate. Because YOLOv5 can be trained on a smaller dataset and still achieve excellent accuracy, researchers and developers with limited data can use it more easily [17], [18]. There are various versions of YOLOv5, with variations in model size and performance: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. The tiniest and fastest variant is YOLOv5s, and the most comprehensive and accurate version is called YOLOv5x [19], [20]. Studies have investigated YOLOv5's performance in various object detection tasks, highlighting its versatility across multiple domains. Research has shown encouraging results in detecting vehicle license plates [21], [22], and in Unmanned Aerial Vehicle (UAV) applications, where YOLOv5 effectively identifies objects from aerial perspectives [23].

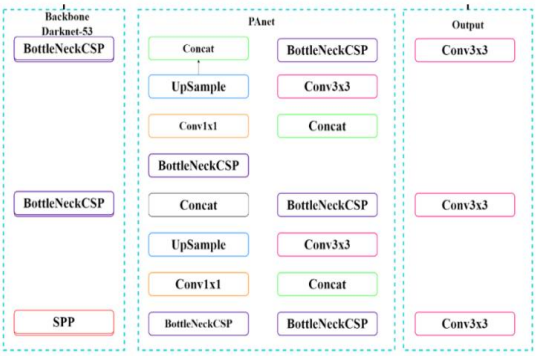
3. METHODOLOGY

- RESEARCH WORKFLOW AND EXPERIMENT SETTING

The experimental design involves partitioning the dataset into two groups: Dataset 1, comprising the original unprocessed images, and Dataset 2, containing the images enhanced using histogram equalization (HE) for image grayscale 8-bit, and Dataset 2, containing the images enhanced using histogram equalization (HE) for images RBG, as outlined in the research workflow diagram (Figure 1a). The YOLOv5s object detection model was subsequently trained on both datasets and performance outcomes were analyzed for the comparative. The YOLOv5 architecture, as depicted in Figure 1(b), comprises four primary components: the input, the backbone, the neck, and the output. The backbone model's key responsibility is to extract salient features from the input image, leveraging fundamental building blocks such as Cross Stage Partial Networks (CSP) and Spatial Pyramid Pooling (SPP) to capture rich and critical attributes. The neck network employs feature pyramid architectures, including the Feature Pyramid Network (FPN) and the Path Aggregation Network (PANet), to enable the development of a robust feature pyramid that facilitates accurate object scaling and generalization.



(a) Research Workfolow



(b) YOLOv5 architecture

Figure 1 an overview of the system (a) Research workflow, (b) YOLOv5 architecture.

Table 1 outlines the training parameter values employed for the YOLOv5s model, including a batch size of 8 and 30 epochs, determined experimentally due to memory constraints. Reporting its inference speed on both CPU and GPU, as well as the number of parameters for an input image size of 256 x 256 pixels. YOLOv5s is the smallest variant of the YOLOv5 family, designed for faster inference and deployment on resource-constrained devices, featuring fewer layers and parameters compared to other variants.

Table 1 YOLO v5 model training parameter values.

Parameters	YOLOv5s
Size Of input Image	256
Epoch	30
Batch	8

4. EXPERIMENT RESULTS

- Datasets

The experiment utilized a dataset of 1145 VIN images obtained from Roboflow [24], a platform that provides tools for preparing and annotating datasets for computer vision model training, including data augmentation, labeling, and dataset management. The dataset was partitioned into training images, 222 validation images, and 116 test and quality images, all with a resolution of 256 x 265 pixels. The dataset comprised samples captured from diverse perspectives within the vehicle, exhibiting significant variability in pictorial conditions, including instances affected by noise, rain, divergent lighting, and differing levels of reflection, as shown in Figure **2.**

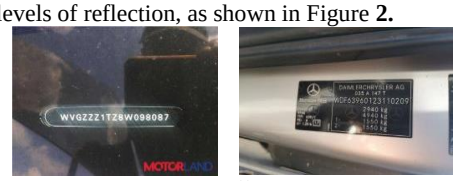


Figure 2 Different VINs

- EVALUATION CRITERIA

The effectiveness of object detection algorithms is typically evaluated using metrics such as precision, recall, average precision (AP), mean average precision (mAP), model parameter count, floating-point operations (FLOPs), and frames per second (FPS)[25]. Average Precision (AP), also known as mean Average Precision (mAP), calculates the average detection precision across varying recall levels, while Intersection over Union (IoU) quantifies the degree of overlap

between predicted and ground truth bounding boxes[26]. The mAP0.5 metric represents the mean average precision at an IoU threshold of 0.5, and mAP0.5:0.95fih represents the mean average precision across a range of IoU thresholds from 0.5 to 0.95, providing a comprehensive assessment of the model's ability to accurately detect and localize objects.

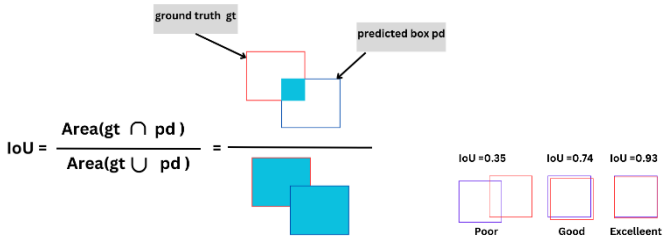


Figure 3 Intersection over Union (IoU) is follow:
a) The IoU is calculated by dividing the intersection of the two boxes by the union of the boxes;
b) Examples of three different IoU values for different box locations.

5. MODEL SELECTION EXPERIMENTS

In this section, present the experiments conducted to evaluate and select the optimal model for object detection using different image preprocessing techniques. In addition to precision, recall, mAP, and F1 score, we analyze Error Rate Reduction, and Frames per Second (FPS) to ensure that the selected model offers a balance between accuracy and computational efficiency.

- Performance Evaluation Using Original Dataset

The model trained on the original dataset serves as a baseline for our experiments. As shown in Table 2, the model achieved a training precision of 97.9% and recall of 95.7%, with a mAP@0.5 of 99.1%. However, the mAP@0.5:0.95 value was relatively low at 62.3%, indicating that the model's generalization ability decreases as the IoU threshold increases.

In the validation and testing stages Tables 3 and 4, similar trends were observed, with the mAP@0.5 hovering around 99.6% and 98.3%, but lower mAP@0.5:0.95 values 66.4% and 66.8%. The confusion matrix for the original dataset (Figure 4) reflects this, showing a small number of misclassifications, with 113 true positives, 5 false positives, and 3 false negatives. This dataset served as a benchmark for evaluating the impact of image preprocessing techniques.

- HE Dataset Performance

When Histogram Equalization (HE) was applied, the model's recall improved slightly, as indicated by 96.3% recall during training and 95.9% recall during validation. However, the precision dropped to 95.1%, likely due to an increase in false positives, as seen in the confusion matrix Figure 5, where 116 true positives and 5 false positives were recorded. Despite the improved recall, the mAP@0.5:0.95 values 61.1%, 65.5%, and 62.9% for training, validation, and testing remained close to those of the original dataset. These results suggest that while HE improves the model's ability to correctly identify true positives, its overall generalization to more stringent IoU thresholds remains limited.

- HE+RGB Dataset Performance

The addition of RGB channels to the Histogram Equalization preprocessing technique (HE+RGB) resulted in mixed performance improvements. While the training recall was relatively high at 94.2%, the confusion matrix Figure 6 shows that the model struggled with a slightly higher number of false negatives and false positives, resulting in 111 true positives, 5 false negatives, and 5 false positives.

The precision remained at 95.1% during training and 95.9% during validation, similar to the HE dataset. However, the generalization capability of the model was somewhat reduced, as indicated by the lower mAP@0.5:0.95 scores of 59.1%, 63.3%, and 62.9% across the datasets. These results suggest that adding the RGB channels did not significantly improve the model's performance compared to HE alone.

- AHE Dataset Performance

The model trained on the AHE dataset showed a marked improvement in recall, precision, and generalization. As illustrated by the confusion matrix Figure 8, the model made only 2 false classifications out of the entire dataset, with 103 true positives and 1 false negative.

The AHE dataset consistently produced strong performance across all evaluation metrics, with precision reaching 98.8% and recall peaking at 99.6% during training. The mAP@0.5:0.95 values were significantly higher than previous datasets, reaching 74.3% during training, 77% during validation, and 73.1% during testing. This dataset exhibited a more robust ability to generalize at higher IoU thresholds, making AHE a highly effective preprocessing method.

- CLAHE Dataset Performance

The CLAHE dataset outperformed all other datasets across every evaluation metric. As demonstrated by the confusion matrix Figure 7, the model produced 104 true positives, with only 2 false negatives and 0 false positives. These minimal misclassifications are reflected in the high precision (P%) and recall (R%) values, which reached 99.4% and 98.6% during training, and 99% and 98.6% during testing. In addition to the high precision and recall, the model trained on the CLAHE dataset achieved the highest mAP@0.5:0.95 values, reaching 75.2%, 77.9%, and 75.2% across training, validation, and testing, respectively. These results indicate that the CLAHE dataset not only improved the detection of true positives but also minimized false positives and false negatives, making it the best-performing dataset across all experiments.

Table 2 shown result training

dataset	P%	R%	mAP@0.5%	mAP0@0.5:0.95 %	F1 score
Original	97.9	95.7	99.1	62.3	97
HE	95.1	95.7	98.1	61.1	96
HE+ RGB	96.3	94.2	97.7	59.1	95
AHE	98.8	99.6	99.5	74.3	99.2
CLAHE	99.4	98.6	99.5	75.2	99.55

Table 3 shown result validation

dataset	P%	R%	mAP@0.5%	mAP0@0.5:0.95 %	F1 score
Original	97.7	95.5	99.6	66.4	97
HE	95.1	96.4	97.6	65.5	96
HE+ RGB	95.9	94.6	96.8	63.3	95
AHE	98.8	99.6	99.5	77	99.2
CLAHE	99.4	99.6	99.5	77.9	99.5

Table 4 shown result test

dataset	P%	R%	mAP@0.5%	mAP0@0.5:0.95 %	F1 score
Original	96.8	96.6	98.3	66.8	97
HE	95.9	1	99.2	69.1	98
HE+ RGB	94.9	95.7	97.4	62.9	95
AHE	97.1	98.1	98.9	73.1	97.6
CLAHE	99	98.6	99.4	75.2	98.8

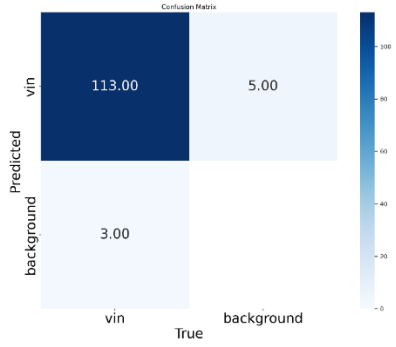


Figure 4 original dataset

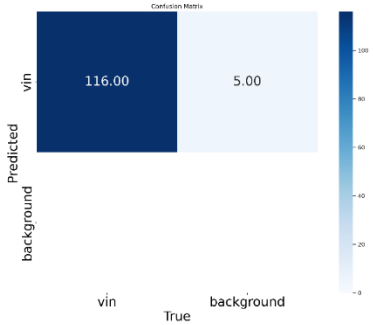


Figure 5 HE+ dataset

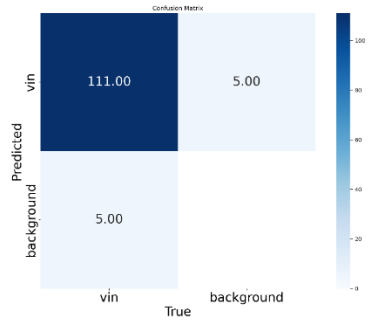


Figure 6 HE+RGB datasets

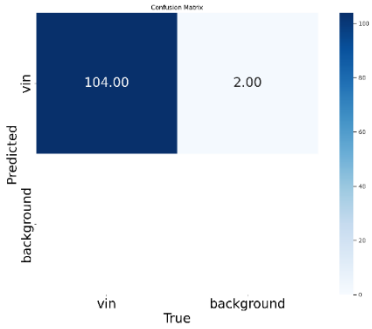


Figure 7 CLAHE dataset

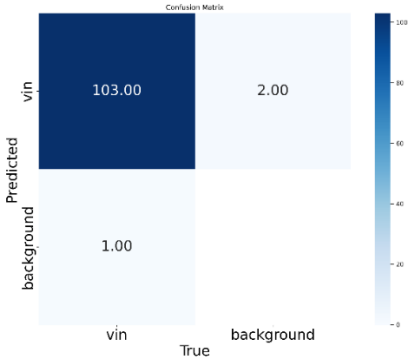


Figure 8 AHE dataset

- Error Rate Reduction

Error Rate Reduction (ERR) is a key metric that quantifies the improvement in prediction accuracy by comparing the error rate of a new model to that of a baseline model (in this case, the model trained on the original dataset). The error rate is the proportion of incorrect predictions (both false positives and false negatives) to the total predictions. The formula for Error Rate Reduction is defined as:

$$ERR = \frac{Error\ Rate_{baseline} - Error\ Rate_{new}}{Error\ Rate_{baseline}} \times 100$$

The Error Rate (ER) itself is computed as:

$$ER = \frac{False\ Positives + False\ Negatives}{Total\ Predictions}$$

Table 5 result Error Rate Reduction

dataset	Error Rate Reduction
Original	Baseline (0%)
HE	37.52%
HE+RGB	-24.96%
AHE	57.19%
CLAHE	71.41%

As shown in the table above, CLAHE achieved the highest error rate reduction of 71.4% compared to the original dataset, highlighting its ability to minimize incorrect classifications. The AHE dataset also showed a significant reduction in error rate at 57.19%, confirming its robustness in improving model predictions. In contrast, the HE and HE+RGB datasets offered more modest error rate reductions.

- Frames Per Second (FPS) Evaluation

Frames Per Second (FPS) measures how quickly the model processes frames, which is essential for deployment in time-sensitive applications.

Table 6 illustrates the FPS performance for each dataset.

dataset	Frames Per Second
Original	3.09
HE	3.08
HE+RGB	210
AHE	3.16
CLAHE	3.20

While CLAHE and AHE datasets demonstrated notable improvements in detection accuracy and mAP values, they also resulted in slightly lower FPS due to the increased computational complexity associated with these preprocessing techniques. Specifically, the CLAHE model achieved an average FPS of 3.02, while the AHE model had an average FPS of 3.28. In comparison, the ORIGINAL dataset maintained the highest FPS at 3.22, with HE and RGB models having average FPS values of 3.26 and 3.22, respectively. The trade-off between FPS and detection accuracy is evident, as more sophisticated preprocessing methods typically incur additional computational overhead. Despite this, the CLAHE dataset provides the best overall balance, offering the highest detection accuracy with only a moderate reduction in FPS. Consequently, for applications where detection performance is crucial, CLAHE is the preferred choice, whereas AHE, HE, and RGB models offer a compromise between speed and accuracy. Based on the results of these experiments, CLAHE remains the optimal choice for image preprocessing in the YOLOv5 object detection model, as it achieved the best results across multiple key metrics, including Error Rate Reduction, IoU Improvement, and mAP@0.5:0.95. Several studies have also corroborated CLAHE's effectiveness in improving object detection results, particularly in applications requiring enhanced contrast, such as medical imaging and autonomous driving [12], [27]. Furthermore, recent research on integrating CLAHE with YOLOv5 demonstrates that this combination yields significant improvements in precision, recall, and error rate, making it highly effective for object detection tasks [14][28]. While it offers slightly lower FPS than other preprocessing methods, its improvements in precision, recall, and error rate make it the superior choice for applications prioritizing detection accuracy. AHE also showed strong performance, particularly in IoU and recall, but did not surpass CLAHE in precision or FPS. Despite offering faster processing speeds, both the HE and HE+RGB datasets lagged behind CLAHE and AHE in terms of error rate reduction and IoU improvement, confirming that more sophisticated preprocessing methods result in better overall performance.

6. Conclusion

This research paper proposes an automated approach to Vehicle Identification Number (VIN) detection using advanced image preprocessing techniques and object detection models. Through the evaluation of various preprocessing methods—including Histogram Equalization (HE), HE+RGB, Adaptive Histogram Equalization (AHE), and CLAHE—using a dataset of 1145 VIN images, the study highlights key findings in optimizing detection performance.

The baseline model trained on the original dataset achieved robust precision and recall but showed limited generalization at higher Intersection over Union IoU thresholds. Techniques like HE and HE+RGB demonstrated improvements in recall but at the cost of reduced precision and generalization capability. In contrast, AHE offered significant advancements, with improved precision and recall and better performance across various IoU thresholds. Nonetheless, CLAHE emerged as the most effective preprocessing method, delivering the highest precision (99.4%), recall (98.6%), and mean average precision (mAP@0.5:0.95) values. CLAHE not only enhanced detection accuracy but also minimized misclassifications, proving to be the optimal choice for VIN recognition. These results underscore the importance of selecting suitable preprocessing techniques to achieve a balance between accuracy and generalization in object detection tasks. The superior performance of CLAHE demonstrates its potential for practical applications in VIN identification and similar fields, offering a robust solution for improving detection accuracy under diverse conditions.

7. References

[1]- J. Celko, “Vehicle Identification Number (VIN),” *Joe Celko’s Data, Meas. Stand. SQL*, pp. 175–178, 2010, doi: 10.1016/B978-0-12-374722-8.00023-2.

[2]- V. A. Adibhatla *et al.*, “Applying deep learning to defect detection in printed circuit boards via a newest model of you-only-look-

once," *Math. Biosci. Eng. 2021 44411*, vol. 18, no. 4, pp. 4411–4428, 2021, doi: 10.3934/MBE.2021223.

[3]- J. Chen, K. Jia, W. Chen, Z. Lv, and R. Zhang, "A real-time and high-precision method for small traffic-signs recognition," *Neural Comput. Appl.*, vol. 34, no. 3, pp. 2233–2245, Feb. 2022, doi: 10.1007/S00521-021-06526-1/METRICS.

[4]- Z. Chen, K. Pawar, M. Ekanayake, C. Pain, S. Zhong, and G. F. Egan, "Deep Learning for Image Enhancement and Correction in Magnetic Resonance Imaging—State-of-the-Art and Challenges," *J. Digit. Imaging 2022 361*, vol. 36, no. 1, pp. 204–230, Nov. 2022, doi: 10.1007/S10278-022-00721-9.

[5]- Y. Jiang, L. Li, J. Zhu, Y. Xue, and H. Ma, "DEANet: Decomposition Enhancement and Adjustment Network for Low-Light Image Enhancement," *Tsinghua Sci. Technol.*, vol. 28, no. 4, pp. 743–753, Aug. 2023, doi: 10.26599/TST.2022.9010047.

[6]- S. Agrawal, R. Panda, P. K. Mishro, and A. Abraham, "A novel joint histogram equalization based image contrast enhancement," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 4, pp. 1172–1182, Apr. 2022, doi: 10.1016/J.JKSUCI.2019.05.010.

[7]- B. S. Rao, "Dynamic Histogram Equalization for contrast enhancement for digital images," *Appl. Soft Comput.*, vol. 89, p. 106114, Apr. 2020, doi: 10.1016/J.ASOC.2020.106114.

[8]- S. Doshvarpassand, X. Wang, and X. Zhao, "Sub-surface metal loss defect detection using cold thermography and dynamic reference reconstruction (DRR)," *https://doi.org/10.1177/1475921721999599*, vol. 21, no. 2, pp. 354–369, Apr. 2021, doi: 10.1177/1475921721999599.

[9]- S. H. Majeed and N. A. M. Isa, "Adaptive Entropy Index Histogram Equalization for Poor Contrast Images," *IEEE Access*, vol. 9, pp. 6402–6437, 2021, doi: 10.1109/ACCESS.2020.3048148.

[10]- S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Comput. Vision, Graph. Image Process.*, vol. 39, no. 3, pp. 355–368, Sep. 1987, doi: 10.1016/S0734-189X(87)80186-X.

[11]- K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," *Graph. Gems*, pp. 474–485, Jan. 1994, doi: 10.1016/B978-0-12-336156-1.50061-6.

[12]- A. Momeni Pour, H. Seyedarabi, S. H. Abbasi Jahromi, and A. Javadzadeh, "Automatic detection and monitoring of diabetic retinopathy using efficient convolutional neural networks and contrast limited adaptive histogram equalization," *IEEE Access*, vol. 8, pp. 136668–136673, 2020, doi: 10.1109/ACCESS.2020.3005044.

[13]- Y. Chang, C. Jung, P. Ke, H. Song, and J. Hwang, "Automatic Contrast-Limited Adaptive Histogram Equalization with Dual Gamma Correction," *IEEE Access*, vol. 6, pp. 11782–11792, Jan. 2018, doi: 10.1109/ACCESS.2018.2797872.

[14]- R. C. Chen, C. Dewi, Y. C. Zhuang, and J. K. Chen, "Contrast Limited Adaptive Histogram Equalization for Recognizing Road Marking at Night Based on Yolo Models," *IEEE Access*, vol. 11, pp. 92926–92942, 2023, doi: 10.1109/ACCESS.2023.3309410.

[15]- J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *ArXiv*, vol. abs/1804.0, 2018, [Online]. Available: https://api.semanticscholar.org/CorpusID:4714433

[16]- J. S. D. R. G. A. F. Redmon, "(YOLO) You Only Look Once: Unified, Real-Time Object Detection," *Cvpr*, vol. 2016-Decem, pp. 779–788, Dec. 2016, doi: 10.1109/CVPR.2016.91.

[17]- K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015, doi: 10.1109/TPAMI.2015.2389824.

[18]- S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path Aggregation Network for Instance Segmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 8759–8768, Dec. 2018, doi: 10.1109/CVPR.2018.00913.

[19]- A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," Apr. 2020, Accessed: Jun. 20, 2024. [Online]. Available: https://arxiv.org/abs/2004.10934v1

[20]- C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 390–391.

[21]- S. Raj, Y. Gupta, and R. Malhotra, "License Plate Recognition System using Yolov5 and CNN," *8th Int. Conf. Adv. Comput. ommun. Syst. ICACCS 2022*, pp. 372–377, 2022, doi: 10.1109/ICACCS54159.2022.9784966.

[22]- K. A, B. D. J, S. P. Gandu, J. Hinn T O, R. Kennedy C, and A. manuel J, "Systematic Number Plate detection using improved YOLOv5 detector," pp. 1–6, Jun. 2023, doi: 10.1109/VITECON58111.2023.10157727.

[23]- P. Sun and X. Ding, "UAV image detection algorithm based on improved YOLOv5," *2022 IEEE 5th Int. Conf. Inf. Syst. Comput. Aided Educ. ICISCAE 2022*, pp. 757–760, 2022, doi: 10.1109/ICISCAE55891.2022.9927693.

[24]- K. Sergeev, "Vin detection Dataset," *Roboflow Universe*. Roboflow, Feb. 2023. [Online]. Available: https://universe.roboflow.com/kirill-sergeev/vin-detection

[25]- T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: challenges, architectural successors, datasets and applications," *Multimed. Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/S11042-022-13644-Y/TABLES/7.

[26]- I. Ahmad *et al.*, "Deep Learning Based Detector YOLOv5 for Identifying Insect Pests," *Appl. Sci. 2022, Vol. 12, Page 10167*, vol. 12, no. 19, p. 10167, Oct. 2022, doi: 10.3390/APP121910167.

[27]- J. P. Q. Tomas, A. G. Carbonell, C. P. A. Aguas, and K. M. Bersabal, "Comparative Analysis of YOLOv7 for Improved Object Detection of Road Markings in Low-Light Conditions: Leveraging CLAHE for Low-Light Image Enhancement," *Proc. 2024 10th Int. Conf. Comput. Artif. Intell.*, pp. 123–128, Apr. 2024, doi: 10.1145/3669754.3669773.

[28]- M. Anas, I. U. Haq, G. Husnain, and S. A. F. Jaffery, "Advancing Breast Cancer Detection: Enhancing YOLOv5 Network for Accurate Classification in Mammogram Images," *IEEE Access*, vol. 12, pp. 16474–16488, 2024, doi: 10.1109/ACCESS.2024.3358686.