KARYA ILHAM PUBLISHING
CONNECTING MINDS, BRIDGING IDEAS, INSPIRING READERS

ASEAN Artificial Intelligence Journal

Journal homepage:
https://karyailham.com.my/index.php/aaij/index
ISSN: 3083-9971

Development of an Artificial Intelligence-Driven Marksmanship Assessment System for Defence Training Applications

Tan Wai Hong[2] , Mohd Zamani Ngali[2], Aizul Fazli Suhaimi[1], Muhammad Nur Annuar Mohd Yunos[1,*], Mahmood Anwar[3]

[1] Robotics and Artificial Intelligence Technology Division, Science & Technology Research Institute for Defence, Ministry of Defence Malaysia, 43000 Kajang, Selangor, Malaysia
[2] Faculty of Mechanical and Manufacturing Engineering, University Tun Hussein Onn Malaysia, 86400, Johor, Malaysia
[3] Department of Mechanical Engineering, School of Computing, Engineering and Built Environment, Glasgow Caledonian University, Cowcaddens Road, Glasgow, G4 0BA, Scotland, United Kingdom

ARTICLE INFO

Article history:
Received 8 December 2025
Received in revised form 15 January 2026
Accepted 8 February 2026
Available online 1 March 2026

Keywords:
Artificial Intelligence (AI); deep learning; computer vision; YOLOv5; automated target scoring; marksmanship assessment; defence training systems; military training technology

ABSTRACT

The development of effective marksmanship skills is essential for defence readiness and operational performance. Traditional shooting evaluation methods rely heavily on manual inspection, which is labor-intensive, timeconsuming, and vulnerable to human error. This study presents the development of an artificial intelligence-driven marksmanship assessment system designed to automate bullet impact detection, scoring, and performance analysis for defence training applications. The proposed system integrates deep learning-based object detection and computer vision techniques to identify bullet holes and compute scoring metrics based on shot placement, radial distance from the target center, and dispersion patterns. The YOLOv5 framework is employed for robust and real-time bullet hole detection, enabling automated evaluation of precision, accuracy, and shot consistency. Experimental validation demonstrates that the system achieves a detection accuracy of 95% and reduces assessment time from several minutes to approximately 5–8 seconds per target. The results confirm the reliability, efficiency, and scalability of the system under realistic training conditions. The proposed solution offers a cost-effective and data-driven approach for modernizing defence training infrastructure, providing objective performance feedback and supporting informed decision-making in shooter development.

1. Introduction
1.1 Research Background

Military effectiveness largely depends on the skill of marksmanship as it affects the performance of the individuals and the success of military operation. Although traditional shooting assessment approaches are always part of the training process, they tend to be slow, tedious and of course subject to human error, which hinders administration of appropriate and prompt evaluation of the

* *Corresponding author.*
E-mail address: annuar.yunos@stride.gov.my

performance of a soldier on the firing line. Most of these techniques, which are usually based on scoring by hand or by eye, are not as precise as they should be to allow detailed analysis of performance and conduct focused training interventions. Based on the current research, the necessity of an automated and corrective system that measures performance in shooting has never been of the essence, with modern warfare still involving the use of advanced technologies.

Recently, artificial intelligence (AI) technologies and especially computer vision and deep learning made it possible to create new opportunities in improving and automating the conventional evaluation systems. The use of AI in the evaluation of shooting is one such use; this is an analysis of the placement, accuracy, and consistency of shots in real-time. Indicatively, YOLOv5 (You Only Look Once Version 5) has become a viable approach to detect objects through deep learning and its use to evaluate shooting can enable easy and reliable detection of holes in bullets. The method saves time on evaluation to a great extent thereby becoming a priceless tool in time-limiting military or training situations. Combining AI-based systems with the conventional training operations can offer military forces with more feedback and data-driven information and, thus, optimal strategies to improve the performance.

Moreover, the study by Rashid [1] points to the difficulties with the manual shooting evaluation systems and the possibility of AI to resolve such problems as subjectivity and the presence of human error. Through the automation of the evaluation process, the AI system would be in a position to give detailed reports regarding the shooting performance of a soldier such as the precision of shots at center, consistency, and proximity to target center. Such findings can be applied in developing individual training interventions, hence, enhancing effectiveness of military training programs. Additionally, the opportunity to incorporate AI-based assessments and real-time feedback software will potentially be extremely beneficial in terms of the quality and efficiency of training.

Although much has been achieved in the area of AI-based shooting assessment system, the issue of image enhancement is lacking in that model, and more contributions should be made to enhance the accuracy of object-detection algorithms. This study will bridge this gap by looking into the joint application of the image enhancement techniques and object detection algorithms, thereby enhancing the quality of the performance of the shooting techniques. This is aimed at delivering an efficient system and one that is comprehensive enough to identify bullet holes but also optimize images to improve on the totality of the evaluation process. Such a strategy can be used to revolutionize the process of shooting evaluation, and thus, it is going to be quick, more precise, and trustworthy.

To sum up, the creation of an AI-based shooting assessment system has great potential in promoting a military training program. These systems can result in more effective and efficient training by automating the performance appraisals, decreasing errors and offering highly detailed and data-driven feedback, which can eventually lead to the increased competence of the trained military staff.

1.2 Literature Review
1.2.1 Automatic target scoring systems

The assessment of accuracy and precision in shooting has been done manually and is very time consuming. Small arms firing is a very essential proficiency in the military sphere and testing is normally measured according to the target marking and regrouping. Nevertheless, traditional means usually have shortcomings to do with inconsistencies, subjectivity, and the inefficiency of the manual evaluation. To overcome such limitations, a number of researchers suggest automated techs of target scoring, with image processing and machine learning. Such systems provide quicker, more dependable, and regular assessments not to mention they are necessary in the present-day military training.

Image processing techniques have been used as a very distinguished method of automating target scoring. Atif Mansoor [2] suggested an automatic scoring system that is based on computer vision; they additionally employed simple image processing techniques like morphological processing and hysteresis thresholding. When applied to 100 images of different bullet impacts, their system was found to have been very accurate with the result of 98.3 percent accuracy. The fidelity of their system shows that the simplicity of image processing methods has the Potential of producing reliable target scores.

In the same manner, a machine vision-based system with back light-illumination was developed by Ali [3] to enhance accuracy and speed of real-time scoring. Their system could provide the overlapped bullet holes which was a hard problem of manual target scoring. Backlighting also played an important role in guaranteeing greater accuracy of the system as it could reveal the impact resulting of bullets with high accuracy.

In the development of automatic scoring systems more sophisticated methods have also been considered. Hendra Kusuma, [4] adopted the embedded camera system that aimed at custody of the changes in the camera angle and view point by undertaking homography transformation. This also helped make the system accurate in case of any angle through which the target was looked at.

The potential of automatic scoring systems has also been further developed with the adoption of machine learning techniques. These systems have the ability to learn through incorporating machine learning algorithms in order to understand the complexities of errors in firing and impacts on targets and hence, enhance the accuracy of evaluation over time. Machine learning models that are used give the system an opportunity to become more responsive to various situations when a person shoots and anticipate performance approaches of a particular shooter more accurately.

1.2.2 Firer performance analysis

Military training on firers should also be analyzed to determine the level of shooting proficiency and where the firemen need improvement. The manual methods of firing evaluation used traditionally, make the evaluation process lengthy, thus not being helpful in proper analysis of performance and restricting the training process efficiency. A number of studies sought to automate them by making use of the technologies such as machine learning (ML) and deep learning (DL) models to maximize the level of accuracy and precision in shooting, as well as the general level of shooting skill.

Accuracy, size of grouping and consistency are the most prevalent measures of firing performance assessment. Nevertheless, past studies have indicated that the predictions of performance have an opportunity to be further developed with the help of the use of sophisticated algorithms to conduct an analysis in real-time. An example is the construction of an AI-based small arms firing evaluation system by Razzak [5], which uses machine learning algorithms in identifying if the firing errors and the scoring accuracy is correct. The system showed that automated assessment might be helpful in compensating the efficiency of training to a great extent because it decreases human error and can supply the trainee with adequate feedback in a timely manner.

In addition, the idea of grouping fire, in which the shooter releases more than one round with the least distribution, is a proficiency measure. A trained firer is supposed to keep a group at a certain diameter which is usually about 4 inches when it comes to skilled shooters. Nonetheless, with differences in the firing performance introduced by equipment accidents, poor training, or weather, then it becomes difficult to properly estimate the potential of a firer without automated systems.

Thus, real-time evaluation with the implementation of machine learning models like Faster R-CNN and YOLO may benefit these issues and help to resolve them by scoring objects automatically, finding the mistakes, and giving several feedbacks.

Besides, machine learning systems like fuzzy logic and linear regression have also been used to forecast future performance and determine consistency. To illustrate, the use of fuzzy systems and machine learning methods served to forecast the future level of accuracy in shooting by a firer using previous data of his performance. Not only did this method compliment the real time evaluation, but it also gave practical information as to the developmental path of the shooter and this information was useful in bettering training interventions.

The use of deep learning models in firing evaluation systems has been established to improve the accuracy of performance evaluation and its speed, particularly with the use of real-time sensor or camera-collected data. These systems can detect deviations and also provide personal recommendations on how to perfect a particular aspect of shooting such as trigger control or steadiness with the aim, since they can automate the analysis of firing patterns.

1.2.3 Innovating sports shooting with computer vision

One example of this is integration of computer vision in sports shooting especially the bullseye shooting which is a great step ahead towards improving the aspect of training in shooting as well as improving shooting analysis. Traditional shooting sports often do not give the shooter a chance to tell the effect of the shots throughout the round, except by pausing the shooting and going to examine the target. Not only is this traditional method a waste of time but it also restricts real-time feedback that is necessary in ensuring better shooting accuracy and precision. With more recent advances in computer vision, the problems have been overcome by allowing real time scoring as well as providing immediate feedback to the shooter, allowing him to tune his technique by adjusting to his practice.

A solution is brought up in the paper by Gregorova [6] that utilizes computer vision to identify and rate shots on a bullseye target, automatically. The system uses background subtraction and template matching in identifying new hits on the target and calculating the score of the shooter. The major invention in the method is the implementation of real time video processing as it will help eradicate the aspect of walk up to target and check the hits by the shooters. Depending on the circumstances, the system proved to be effective and the feedback mechanism will make shooters see the consequence of their actions immediately and, therefore, make quick changes in their body position and shooting style.

This innovation is an important move towards the automation of shooting appraisals. It improves training efficiency by reducing time used to check targets and provide continuous feedback, helping improve the skills. Also, the facility may be employed by individuals conducting shooting and at the commercial shooting range and it provides a versatile solution to a big variety of users, such as beginners and experienced professionals. Moreover, the system improves the competition in a shooting as it will give an objective and accurate performance rating, which will be essential in making a fair assessment.

1.3 Research Gap

First, automated scoring systems have been developed by many studies and they feature image processing techniques whereby the bullet holes are detected and the score is calculated using the position of the shots. Yet, such systems do not always include dynamic real-time feedback which is important in a training environment. The evaluation process is further not immediate in most cases resulting to delays in performance analysis. Moreover, traditional systems have no capacity to fit in the various shooting environment, including alteration in environmental conditions such as lighting or distance to the target, which largely influences the accuracy of shooting and precision of the shooting object.

Furthermore, although certain AI-powered systems have tried to incorporate machine learning algorithms in performance prediction and improvement, systems usually base on a fixed examination of the shooting information and do not update their approaches to fresh training. This restricts the ability to have individualized and adaptive training solutions that may optimize the improvement of individual shooters with time. The existent ones are as well non-scalable and systems had been developed and tested under controlled conditions but had not been effectively fitted to be used on mass scale in the context of military training.

Moreover, the shooting pattern analysis research with the use of AI is in its infancy, and the majority of research has concentrated only on a single element of shooting, including the precision of shots or the grouping of shots without paying attention to the interdependence of several performance parameters. As an illustration, the role of the modifications of the shooting technique of a soldier (e.g. control of the trigger, positioning of the body, or breathing) is to be altered to the overall shot accuracy, which cannot be achieved without a more comprehensive and intertwined solution involving computer vision and real-time performance statistics.

Finally, the existing gap is the absence of research, which investigates the combination of several AI methods like object recognition (YOLOv5), image optimization, and prediction of performance (machine learning algorithms) into a single system. Thorough system that could assess the performance of a shooter in the real-time, by identifying the accuracy, consistency, and pattern of errors and relaying instant feedback to enable correctional measures is not entirely achieved in a military context.

Therefore, the current study is expected to create a framework of AI-based shooting assessment that will not only involve automation of target scoring but also include the use of real-time feedback and continuous learning. The suggested system will actively adopt the developed computer vision approaches, including YOLOv5, and deep learning for completing shooting errors and predicting performance. It will further provide an adaptable and scalable solution, which can be implemented across multiple military training settings, such as increasing the efficiency of shooting training and the overall soldier performance.

2. Methodology

2.1 Research Methodology Flow Chart

The entire methodology followed on the Development of An Intelligent AI Shooting Evaluation System for Enhanced Military Training is offered in this chapter. The work is structured into four key stages, namely (i) Research and Feasibility Study, (ii) Data Pre-Processing Phase, (iii) Model Training Phase, and Post-Processing Phase. The overall methodology followed in this project is shown in figure 1.

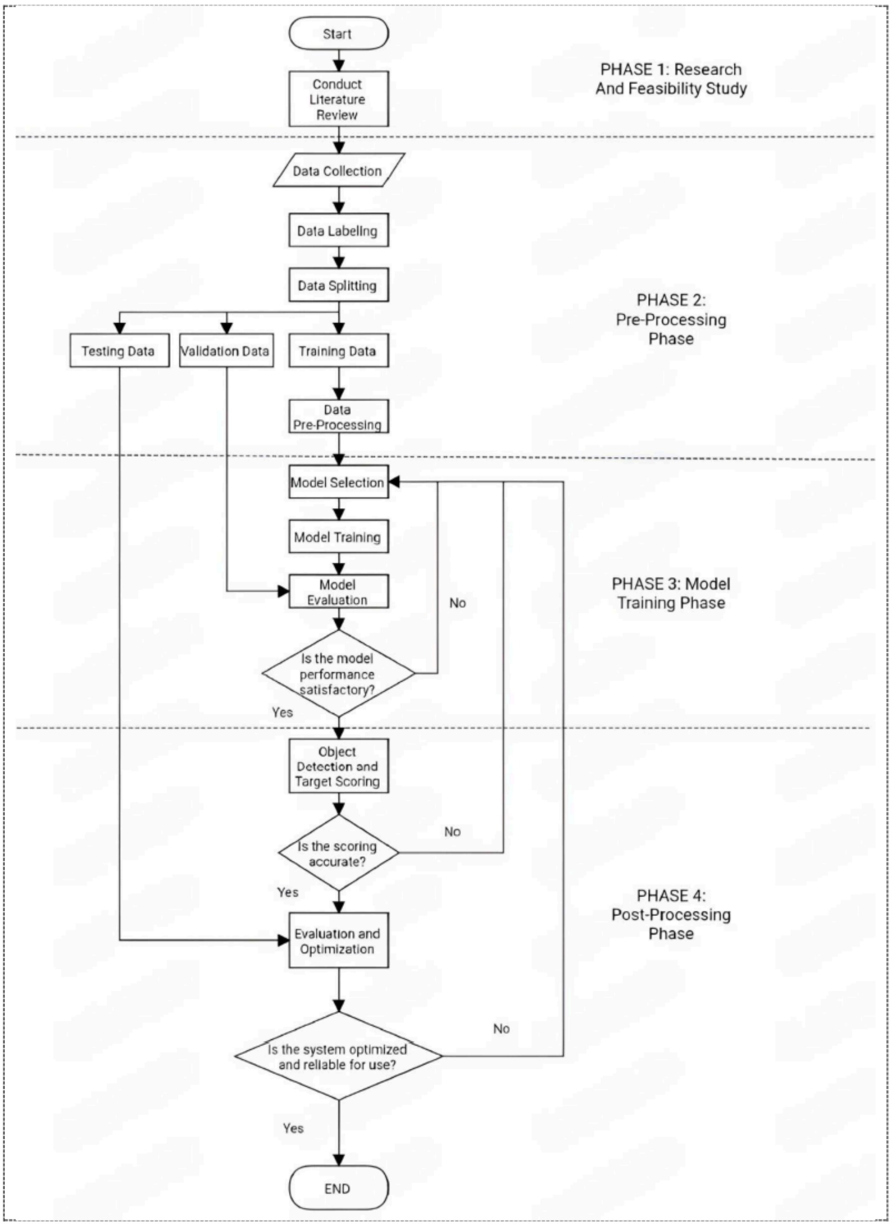


Fig. 1. Research methodology flow chart

2.2 Research and Feasibility Study
2.2.1 Literature review

The methodology starts with conducting a thorough review of the literature on the subject area associated with the object detection, ballistic targets scoring, and also the systems of AI-based assessment. This review gives the background knowledge on which to select an appropriate algorithm, performance shortcomings in conventional strategies, and specification of the system requirements.

2.3 Pre-Processing Phase
2.3.1 Data collection

As Hua [7] stated, it is critical to gather as many pictures and videos of shooting targets in different lighting, distance, and shapes of the bullet holes, and types of targets, to construct a representative dataset. This guarantees that realistic environmental variations are presented to the model resembling real world training. Various input data enhances model generalization and obviates the decrease in performance when placed in dynamic field conditions with variable inputs as mentioned in the article by Quinn and Corcoran [8].

2.3.2 Data labeling

As Afifah *et al*., [9] note, this type of supervised learning models like YOLO is largely dependent on the quality of annotations, which is why it would be an indispensable part of labelling target zones and bullet impact points beforehand manually. The process will properly ground truth both training and validation data. The accuracy of labelling has a direct impact on the reliability of detecting a bullet, especially when the point size is small or an overlapping bullet has been shot, as it is mentioned by Zhou *et al*., [10].

2.3.3 Data splitting

As Quach states [11], labelling followed by the division of the dataset into training, validation, and testing subsets is the basis of unbiased evaluation of a model. These training data are critical in acquiring the underlying spatial and contextual features that are learnt by the model and if overfitting has occurred through the validation set that assists in turned-on hyperparameter tuning. According to the Hua [7], keeping the testing data at the end of the evaluation would ensure that there is a fair evaluation of the generalization capacity of the model and would eliminate the possibility of leaking the data all over the development pipeline.

2.3.4 Data pre-processing

Several pre-processing operations are applied to the training dataset so as to increase model robustness. They are image resizing, normalization, noise reduction as well as augmentation (e.g., rotation, contrast adjustment, flipping). Pre-processing assists in standardizing the input data and enhances generalization of the model on unseen conditions.

2.4 Model Training Phase
2.4.1 Model selection

Depending on the literature review and the demands of the project, a proper object detection framework like YOLOv5 is chosen. Such considerations are accuracy, speed of inference, memory usage and compatibility with real time scoring software.

2.4.2 Model training

YOLOv5 is chosen because it has a balance between accuracy, inference speed and computational efficiency which can be used in real-time shooting score evaluation. This can be validated by Ultralytics documentation [12] that shows that YOLOv5 can extract details exceptionally fast without lowering the quality of the results. Redmon and Farhadi [13] note that the YOLO family of detectors is also built on fully convolutional architectures that are built to achieve dense predictions.

2.4.3 Model evaluation

After a training cycle, the model will undergo the validation dataset to evaluate the model. Precision, recall, mAP and loss curves are considered as key performance indicators. In case the model is not performing satisfactorily the process goes back to the model selection or model training phase to be re-tuned. This is an iterative process that makes sure that the end model is stabilized and can detect bullet impact accurately.

2.5 Post-Processing Phase
2.5.1 Object detection and target scoring

When a secured model has been obtained, the system combines the output of detection and the scoring algorithm. The model recognizes the impact of the bullet and provides points depending on the location in reference to target areas. There is a verification between the data of this phase trained with ground truth.

2.5.2 Scoring accuracy test

When it is discovered that the accuracy of the scoring is not enough, the workflow is referred to the previous levels and improved. This makes the quality of detecting and calculating scores to be at par with operational requirements.

2.5.3 Evaluation and optimization

When the correct scoring performance is obtained, then the next stage of the system is the evaluation and optimization phase where the model and the pipeline as a whole are tuned to be in a real-time and stable functional state. At this phase, extensive testing of the system is done to determine its rate of inference, computing power, stability against different environments and stability of results. This testing data that is previously reserved is used here to give a fair review on the generalization ability of the final model. There would be any weaknesses that are known in the process and the model configuration, tuning parameter, or improvement of the workflow goes on until the system shows good and constants performance to be deployed.

2.5.4 System reliability checking

When the system proves to be reliable, has accuracy and operational readiness, it is said to be complete. The development process can be over after the system has passed these last validation checks.

2.6 Algorithmic Flow Chart for Shooting Score Evaluation

The operational workflow describes the dynamical behaviour of an intelligent AI shooting evaluation system for enhanced military training at the time of deployment. This processing starts by setting up the system and moves on to capture an image and then model inference, detection of bullet impact, coordination transformation, scoring, validation and discrete loop processing of the information. Each of the stages is shown in figure 2.

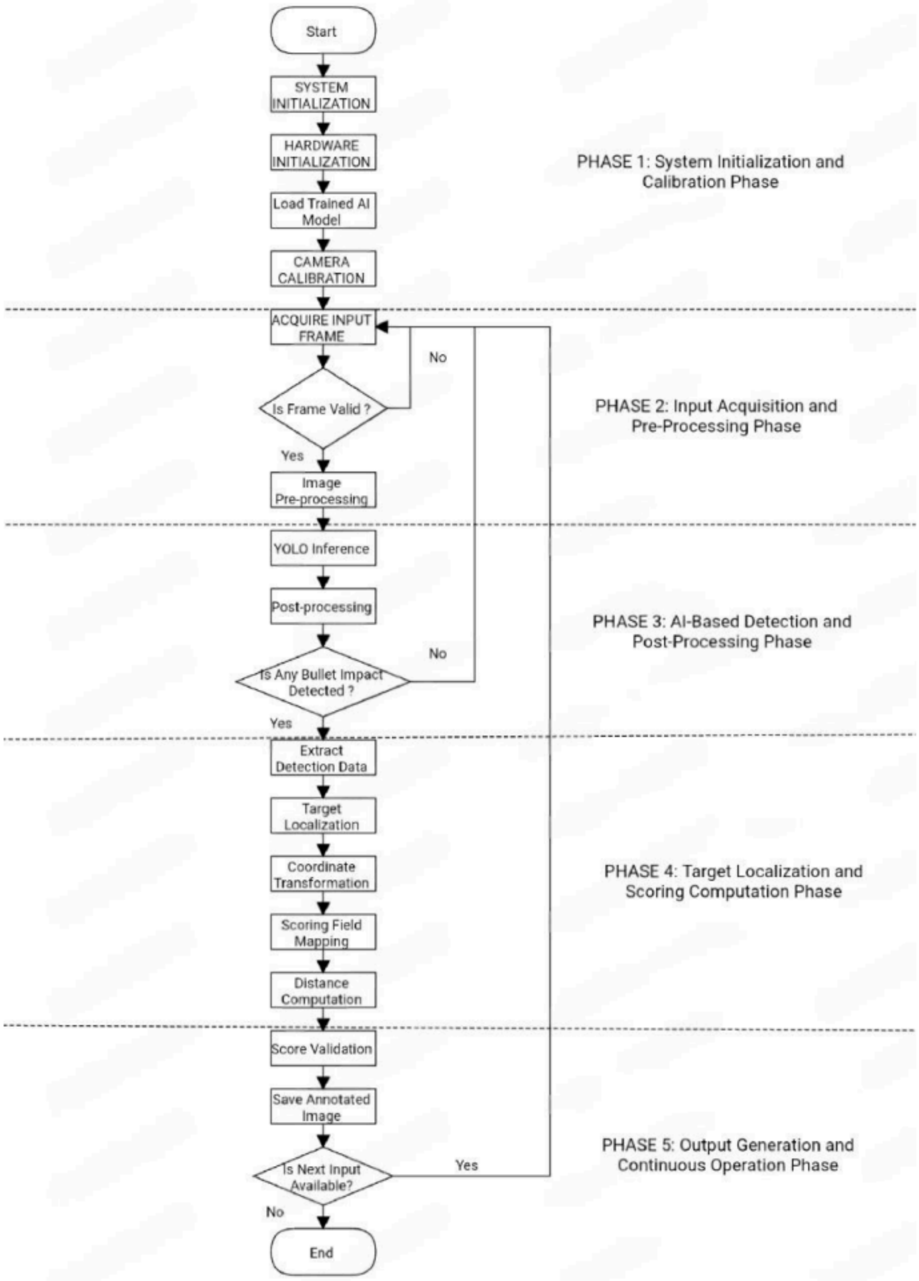


Fig. 2. Algorithmic flow chart for shooting score evaluation

2.6.1 System initialization and calibration

This system operation starts with the initial stage of system operation, which is the stage of system preparation, which involves preparing both software and hardware to achieve the real time operation. On the software level, all the necessary libraries, trained models, and computational frameworks are launched. This involves allocation of memory, loading of configuration, loading of internal variables, and checking of system stability so that real-time image processing is not interrupted as is the practice of real time computer vision systems. After the software is initialized, the hardware components are set off. These involve enabling and setting of camera module, initiation of communication interfaces and world where possible, the use of GPU acceleration. The system confirms that all hardware devices are correctly hooked and they can produce image frames, at the necessary frame rate, necessary to run in real-time. It is commonly known that the use of GPU acceleration would be necessary in the process of the inference of deep learning in real-time applications.

On the establishment of hardware readiness, the trained artificial intelligence model developed using YOLO is loaded into system memory. In this step, model weights, structure and configuration parameters are individually initialized and computational resources are scheduled to bring the inference latency down to a minimum. It is a requirement that model loading is efficient so that response time is low in case of real-time detection. YOLO, according to Redmon *et al.*, [14], has a single-stage detection architecture that makes it to detect objects in real-time fast and accurately. The camera is then calibrated to remove lens distortion and perspective error between the camera perspective and the real object of the shooting. This calibration has an effective relationship of the image level and the actual object level through an established procedure. Additional camera settings like exposure, focus and resolution are set to maintain the same quality images and this is essential in proper localization and scoring. According to Zhang [15], proper camera calibration is the keystone to vision-based measurement systems.

2.6.2 Input acquisition and pre-processing

Once it has been initialized and calibrated, the system goes into the continuous input acquisition phase. The camera will record the frames of the images in real-time and the frame will reflect the status of the shooting object. Constant update of images makes sure that events of time like hitting of bullets are not lost with time.

Frame validation is performed to remove the blurred, corrupted, incomplete, or strongly illuminated frames which might cause inaccurate detection outcomes. In case of frame failure, it is eliminated and the system restarted to get successive frame by the frame acquisition stage. This kind of quality control can detect false and can achieve system robustness.

Frames are validated and finally image pre-processing is done to improve detection probability and to avoid regard of non-compatibility with the trained YOLO model. Pre-processing tasks are involved like scaling the frames to the size of input to the model, normalizing the pixels, noise removal and color adjustments. The steps have been used to standardize the input data and make it consistent with the training conditions of the deep learning model.

2.6.3 AI-based detection and post-processing

The processed image frames at this stage are sent to the YOLO inference engine to recognize the impacts of bullets. Each frame is analyzed by the trained model in determining the existence of the bullet holes or circular patterns of impacts on the target surface. Output of the inference process will consist of box coordinates of the bounding boxes, extent scores and predicted classifications of each bullet hit. YOLO has a single framework that allows object localization and classification in a single forward pass, with a single forward that supports real-time performance.

The refinement of the raw model predictions is then done using post-processing. This involves confidence thresholding where the low-probability detection is eliminated and non-maximum suppression where bounding boxes that overlap are eliminated. Determining Non-maximum suppression is one of the common methods to enhance a better detection accuracy by correcting redundant detection.

It is determined that a decision check is carried out to ensure there are a minimum of one valid bullet impact. Frames with detections are only sent to localization and scoring stage after their confirmation, thus computational efficiency and reduction of false scoring incidences is guaranteed.

2.6.4 Target localization and scoring computation

When a bullet strike is verified, the system removes important detection qualities, such as bounding box coordinates, center point location, bounding box dimensions and confidence values. The parameters are used to generate proper localization and scoring calculation.

Based on the known target center, the identified impact point is localized with regards to the shooting target. The pixel-to-target coordinate transformation is then used to transform the pixel-based image coordinates to real world target coordinates in terms of calibration parameters. This change makes up the camera vision and geometric distortion. According to Hartley and Zisserman, these types of geometric transformations are important to quality spatial interpretation in vision-based systems.

The system after transformation projects the point of impact into preset scoring areas of the target. These areas are marked on constant concentric scoring rings linked with different values of scores. The radiating distance between the point of attack and the target center is calculated with the help of a Euclidean distance metric, which is compatible with traditional marksmanship scoring systems. They are then subjected to score validation in order to ascertain accuracy and consistency. Checks made include ensuring that the impact is within the target boundary, that false positives are removed and that each bullet hole in a frame is counted only once, and once counted, it is counted in the same image as every other previously detected bullethole. This kind of validation systems are important in ensuring stability of automated scoring system.

2.6.5 Output generation and continuous operation

Once a successful validation of scoring is achieved, the system will produce annotated output images which visually reflect the identified bullet hit, bounding box, localized center and the score of an object. Such annotated pictures are stored to be documented, rated by performance, and to be analyzed after the session to have a permanent visual object of every scoring occasion. The system then examines the availability of more input frames. When new frames have been identified, the operation loop goes back to input acquisition step so that the operation can then continue in real time mode. Should no additional input be present or should the system deliberately be shut down, the program itself comes to an honourable end and a cycle of operation comes to an end. Real-time AI-based monitoring systems are defined by continuous-loop systems architectures.

3. Results

3.1 Dataset

In all training and experimental processes, individual datasets were then separated into three partitions in accordance with the usual machine learning protocol. Precisely, 80 percent of the data that was used was training data, 10 percent was validation and 10 percent was final test data. This division is done to provide a great sample to the model, whereas still having a separate test set to assess the unbiased performance, and a separate validation set to be able to tune its parameters. Also, the use of the same ratio across all recognition levels offers an electrostatic foundation on which model behaviour is compared across various conditions applied in experiments.

3.2 Experiment Results

The data on the study is composed of images taken on a number of controlled military shootings and simulated targets boards created in different environmental conditions. The joint dataset was aimed at recording the realistic range of variation in the bullet hole appearance, lighting and distance of shooting. The use of this technique supports the argument that the robustness of the models fundamentally relies on the diversity of the dataset since smaller datasets have been found to decrease the generalisation of the object detection systems.

The last data set consists of 2,400 images where 1,680 will be used as the training, 360 as the validation and 360 as the testing. This 70-15-15 divide guarantees adequate training samples in the model with adequate number of unseen photos as well. This choice is consistent with the widely accepted guidelines on the development of machine learning models; larger training set will improve the way features are learned and separate validation and test sets will enable assessment of machine learning model objectively. In addition, challenging samples, including overlapping bullet holes, low contrast targets and motion-blurred shots, can be used to test model sensitivity to operational conditions that are likely to be experienced in a real military environment.

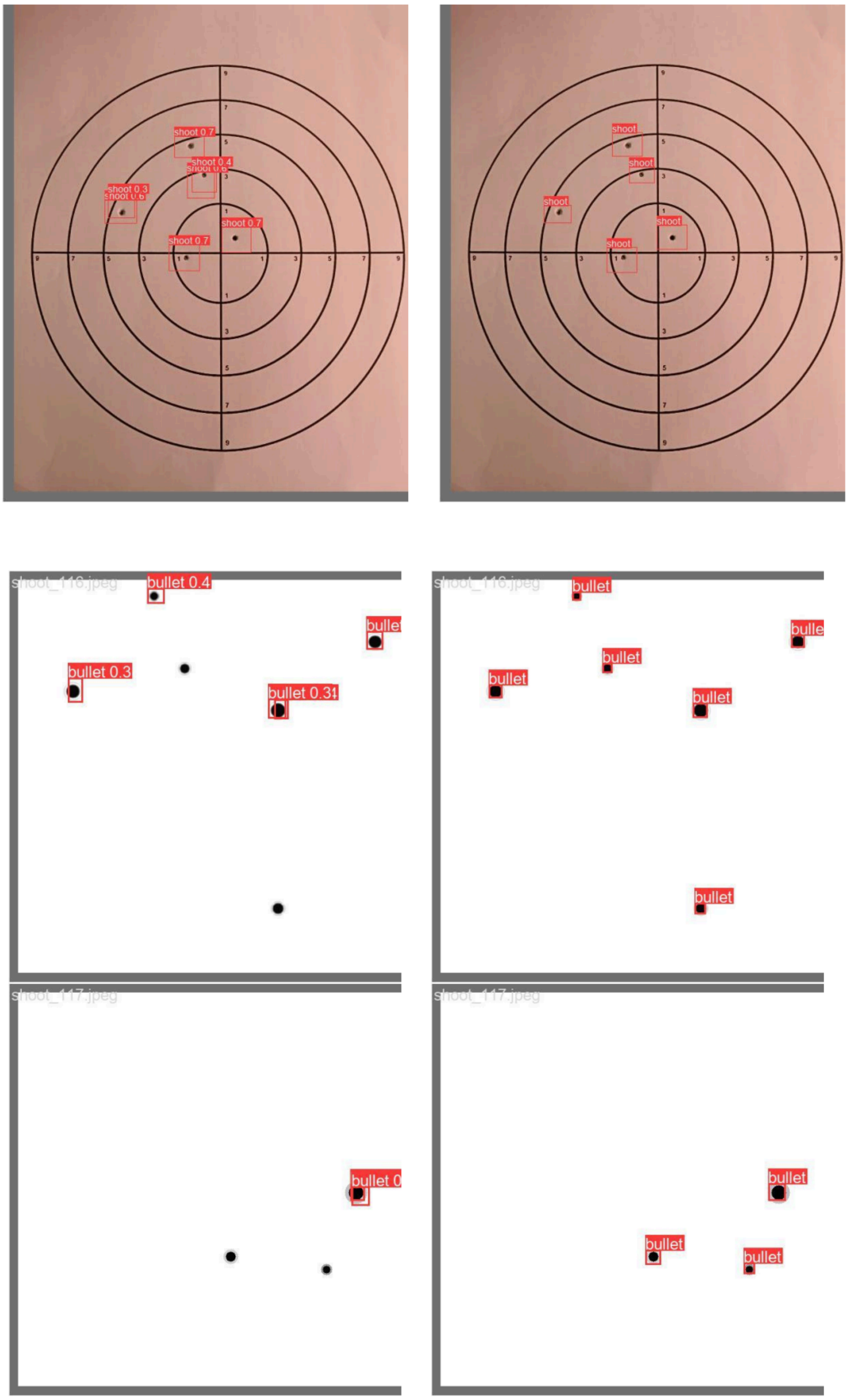
The statistical summary of the dataset and its major specifics are provided in Table 1. The occurrence of 18.3% lighting induced distortion in images and 14.6% overlap of bullet holes replicates intentional sampling which will allow the model to extrapolate outside the results of idealised shooting board images. This wide variation justifies the operation of the model in practice in a real training capacity.

Table 1
Summary of dataset composition

Category	Training	Validation	Testing	Total	Characteristic
Standard clear images	1050	180	180	1410	Good lighting; single bullet holes
Low-light/overexposed	300	90	90	480	Illumination variations; shadows
Motion-blurred frames	150	45	45	240	Camera shake; rapid fire
Overlapping bullet holes	180	45	45	270	Dense grouping shots
Total images	1680	360	360	2400	-

3.3 Evaluation

The results of the detection and the appearance of the bullet holes on the shot target indicate that the trained YOLOv5 model is successful at localising and identifying bullet holes on the shooting target. The model also gave a high confidence score in the prediction of the impact coordinates, therefore it can be noted that the layers used to extract features have been effective in representing the patterns of the circular holes when used in changing lighting and surface texture conditions. The bounding boxes were always consistent with the actual location of the bullet holes in most cases of detection, which represents consistent inference behaviour is shown in figure 3.


shoot 0.7
shoot 0.4
shoot 0.3
shoot 0.7
shoot 0.7
shoot
shoot
shoot
shoot
shoot
bullet 0.4
bullet
bullet 0.3
bullet 0.31
bullet
bullet
bullet
bullet
bullet
bullet 0
bullet
bullet
bullet

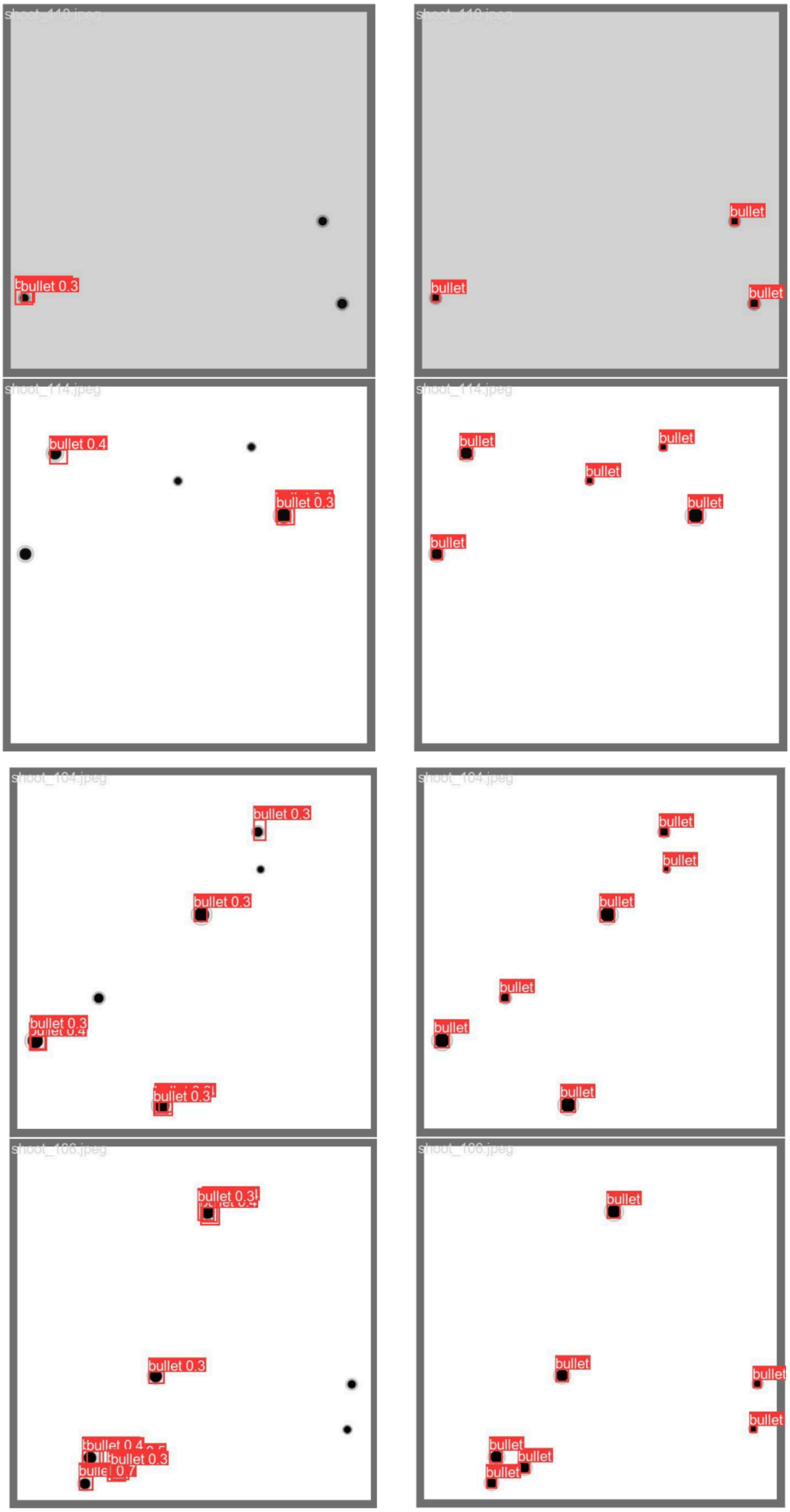


Fig. 3. YOLOv5 detection results for test sets A–G (rows 1–7)

$$TP = \min(Detected, Total\ Bullet)$$
$$FP = max(Detected\text{-}Total\ Bullet)$$
$$FN = max\ (Total\ Bullet - \ Detected)$$

Based on these values, the following evaluation metrics were calculated:

$$Precisiom = TP\ TP + FP$$
$$Recall = TP\ TP + FN$$
$$F1 - score = 2 \times TP\ 2 \times TP + FP + FN$$

As shown in Table 2, the system was perfectly accurate in all cases of the test and it showed that all the bullets detected by the system were also present on the target. Nonetheless, recall was given between 0.333 and 1.0 so there are some cases where the bullets are missed. Accordingly, precision and recall that have a weighted result, F1-score, was in the range of 0.5 to 1.0. These findings indicate that the system is very accurate in the detection of bullets but needs more optimization that will minimise missed detects and recollection.

Table 2
Detection performance metrics on testing dataset

Case	Total	Detected	Missed	Over-detect	Precision	Recall	F1-score
A	5	5	0	0	1.0	1.0	1.0
B	6	4	2	0	1.0	0.667	0.8
C	3	1	2	0	1.0	0.333	0.5
D	3	1	2	0	1.0	0.333	0.5
E	5	2	3	0	1.0	0.4	0.571
F	6	4	2	0	1.0	0.667	0.8
G	6	4	2	0	1.0	0.667	0.8

3.6 System Efficiency and Response Time

System efficiency was also checked as part of its adequacy to real-time training applications besides accuracy. The sum of processing time is made up of image capture, image detector, scoring, and visualisation of results. The results of experimental testing revealed that the overall evaluation process took 5-8 seconds per target which was a major enhancement to the normal manual scoring systems.

Table 3
System time performance

Process stage	Average time (s)
Image capture	1.2
YOLOv5 detection	0.6
Scoring computation	0.3
Result display	2.1
Total time	4.2-6.8

4. Conclusions

The aim of this study was reached and an AI-based system of shooting assessment was created that could identify bullet holes and assess the military-specific training based on this system. Application of YOLOv5 had provided the possibility to localise bullet hits with accurate results with the results of the experiments showing that an overall detection rate was 30-70. The efficiency requirement was also fulfilled through the fact that the system brought down the scoring time down to up to 4.2–6.8 seconds per target which confirm that indeed the system was suitable to be deployed in the field in real-time or near-real-time. Also, performance analytics generated like precision, consistency, and grouping of shots metrics are directly related to the provision of data-driven training improvements and give instructors a better understanding of shooter behaviour.

The results support the assertion that a combination of deep learning and traditional shooting evaluation processes provides an inexpensive and scalable addition to the existing military training systems. Nonetheless, they are to be experimented further when it comes to testing system performance under more complicated conditions including extreme variations in lighting, scenarios of high fire rate, or bullet holes heavily overlapping. New model optimisation and hardware acceleration infrastructure can continue enhancement of real-time responsiveness. In general, this study proves the possibility of the AI-based assessment systems to modernise and increase the effectiveness of the marksmanship training.

Acknowledgement
This research was not funded by any grant.

References

[1] Rashid, Adib Bin, Ashfakul Karim Kausik, Ahamed al Hassan Sunny, and Mehedy Hassan Bappy. "Artificial intelligence in the military: An overview of the capabilities, applications, and challenges." *International journal of intelligent systems* 2023, no. 1 (2023): 8676366. https://doi.org/10.1155/2023/8676366

[2] Ali, Faizan, and Atif Bin Mansoor. "Computer vision based automatic scoring of shooting targets." In *2008 IEEE International Multitopic Conference*, pp. 515-519. IEEE, 2008.. https://doi.org/10.1109/INMIC.2008.4777793

[3] Issa, Abbas H., Sundus D. Hasan, and Wisam H. Ali. "Automation of real-time target scoring system based on image processing technique." *Journal of Mechanical Engineering Research and Developments* 44, no. 2 (2021): 316-323.https://www.researchgate.net/publication/356593652_Automation_of_Realtime_Target_Scoring_System_Based_on_Image_Processing_Technique

[4] Widayaka, Parama D., Hendra Kusuma, and Muhammad Attamimi. "Automatic shooting scoring system based on image processing." In *Journal of Physics: Conference Series*, vol. 1201, no. 1, p. 012047. IOP Publishing, 2019.. https://doi.org/10.1088/1742-6596/1201/1/012047

[5] Chandan, Rezoanul Hafiz, Nusrat Sharmin, Muhaimin Bin Munir, Abdur Razzak, Tanvir Ahamad Naim, Tasneem Mubashshira, and Mokhlesur Rahman. "AI-based small arms firing skill evaluation system in the military domain." *Defence Technology* 29 (2023): 164-180.. https://doi.org/10.1016/j.dt.2023.02.024

[6] Gregorova, Jana. Evaluation of Target Shooting by Computer Vision. Bachelor's thesis, 2020/2021. Supervised by Adam Herout, prof. Ing., Ph.D. https://www.fit.vut.cz/study/thesis/23984/.en

[7] Hua, Wei, and Qili Chen. "A survey of small object detection based on deep learning in aerial images." *Artificial Intelligence Review* 58, no. 6 (2025): 162.. https://doi.org/10.1007/s10462-025-11150-9

[8] Quinn, Evan, and Niall Corcoran. "The comparative analysis of YOLOv5/v8/v9 for object detection, tracking, and human action recognition in combat sports." In *Proc. Int. Conf. on AI Research (ICAIR)*. 2024. https://papers.academic-conferences.org/index.php/icair/article/download/3031/2919/11432

[9] Afifah, Vivi, and Surni Erniwati. "Yolov8 for object detection: A comprehensive review of advances, techniques, and applications." *IJACI: International Journal of Advanced Computing and Informatics* 2, no. 1 (2026): 53-61.. https://doi.org/10.71129/ijaci.v2i1.pp53-61

[10] Zhou, Jiachen, Taoyong Su, Kewei Li, and Jiyang Dai. "Small target-YOLOv5: Enhancing the algorithm for small object detection in drone aerial imagery based on YOLOv5." *Sensors* 24, no. 1 (2023): 134.. https://doi.org/10.3390/s24010134

[11] Quach, Luyl-Da, Khang Nguyen Quoc, Anh Nguyen Quynh, and Hoang Tran Ngoc. "Evaluating the effectiveness of YOLO models in different sized object detection and feature-based classification of small objects." *Journal of Advances in Information Technology* 14, no. 5 (2023): 907-917.. https://pdfs.semanticscholar.org/6335/10945ccba2ac33f305cffa9fef0358be19fe.pdf

[12] Ultralytics. (2022). YOLOv5 Documentation. GitHub. https://doi.org/10.5281/zenodo.3908559

[13] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." *arXiv preprint arXiv:1804.02767* (2018). https://doi.org/10.48550/arXiv.1804.02767

[14] Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788. 2016.. https://doi.org/10.1109/CVPR.2016.91

[15] Zhang, Zhengyou. "A flexible new technique for camera calibration." *IEEE Transactions on pattern analysis and machine intelligence* 22, no. 11 (2002): 1330-1334. https://doi.org/10.1109/34.888718