



Real Time Classification of Retail Theft Utilizing YOLO Algorithm

Ali Salah Hameed^{*}, Taha Mohammed Hasan, Rokan Khaji

Computer Science Department, College of Science, University of Diyala, Baquba 32001, Iraq

Corresponding Author Email: scicomphd232402@uodiyala.edu.iq

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.300610>

ABSTRACT

Received: 6 May 2025

Revised: 8 June 2025

Accepted: 22 June 2025

Available online: 30 June 2025

Keywords:

retail theft detection, YOLOv8, real-time surveillance, computer vision, object detection, anomaly behavior classification, machine learning

With the rapid advancement of computer vision technologies, human behavior detection systems in surveillance environments have become a vital research area, especially in security applications such as shoplifting surveillance. This study aims to develop a classification model based on still images to detect suspicious behavior in a shopping environment. Current theft detection systems struggle with real-time processing; our YOLOv8-based framework addresses this by achieving 95% accuracy with low latency (12 ms per frame). This makes our approach suitable for being integrated in real-time monitoring systems where it guarantees an early and robust detection of abnormal behavior. Comprehensive comparison with other state-of-the-arts also confirms the superiority of the proposed method in terms of speed and detection accuracy. The model was trained using the UCF Crime dataset, as well as manually collected suspicious images from multiple sources. The categories comprised (normal behavior) and (suspicious behavior). The model was trained for 150 epochs and the model parameters were fine-tuned to obtain the best performance. Different performance metrics including precision, recall, F1-score, confusion matrix analysis, and visual results of the output of the model, were assessed. This paper is a first step in the development of an intelligent system for detecting suspicious in-store behavior. This system will be extended to analyze temporal behavior with video sequences, with a view to providing a richer and more accurate understanding of theft pattern in its temporal context.

1. INTRODUCTION

Shoplifting, also known as retail theft, poses a significant problem for businesses and owners, resulting in security challenges and economic losses [1]. This theft takes various forms, such as concealing an item under clothing, placing it in a personal bag, or concealing it in a pocket [2]. The primary function of camera systems within shopping environments is to serve as a preventative and theft-reducing tool. However, despite their ubiquity and availability, the number of shoplifters is on the rise, and the losses incurred by store owners are also increasing, with the number of shoplifters reaching approximately 27 million, and the losses also reaching \$13 billion annually [3]. The reason for the increase in these thefts is the limited human capacity for continuous monitoring. Consequently, it is imperative to adopt an intelligent monitoring system that classifies theft incidents and analyzes individual behavior in real time [4]. Researchers have focused their attention and efforts on developing applications as a result of advances in image processing and computer vision, which have opened the door for them to address the challenges associated with developing intelligent systems in a practical and effective manner, particularly in the field of human behavior analysis [5]. This study, which used YOLO v8, seeks to automatically classify shopper behavior within shopping environments, as this will detect suspicious activity and enhance security measures early [6]. Real-time object

detectors should follow an optimal trade-off between speed and accuracy. Towards this end, many works have been put forward to design efficient and effective architectures for real-time object detection [7]. Yolo version is based on a deep neural network that has been optimized to increase learning efficiency and reduce information loss across layers [8]. The image entering this model is divided into a grid. Each set of predictive boxes may identify a specific object within this grid, and each object has a probability [9]. A combination of mathematical, logical, and nonlinear steps is employed, as well as procedures to find optimal solutions, known as optimization techniques. These techniques strike a balance between correct detection and minimizing losses [10]. Accordingly, YOLOv8 will be an important tool for operating intelligent systems to monitor suspicious behavior in shopping environments, including retail stores [11]. The YOLO architecture transforms object detection by hierarchically synthesizing a single neural network that collectively performs bounding box regression, and simultaneously predicts object types contained in the images [12]. Unlike traditional approaches that separate the proposal and classification phases, YOLO offers a complete solution implemented in a single step, reducing processing time and increasing efficiency [13]. It is therefore highly suitable for real-time applications such as smart surveillance. Newer structural optimization methods are added in its latest version YOLOv8 whereby both the accuracy and speed can be improved [14]. According to the above, the YOLOv8 model

was selected as the backbone detector of the proposed system for its high real-time detection and classification efficiency, as well as its balance between accuracy and inference velocity [15]. This model fits perfectly with the manually labelled data used in this work, and is used to localize people within images, and classification between normal and suspicious behavior. When the advantages of YOLOv8 (such as network architecture optimization) were utilized, the system can effectively capture fine-grained visual characteristics. Further, the model can be extended to the temporal analysis models, such as GRU or LSTM, due to its flexible architecture. This will be evaluated in the next study, which will also concentrate on the analysis of the temporal sequence of observed behaviors improving the system's accuracy by considering complex behaviors, such as product concealing and movement hesitating. For this study, the YOLOv8 model was trained on the UCF Crime Database as well as an ad-hoc set of images manually collected from various sources representing abnormal behavior within the shopping scene. The model was trained across 150 epochs with the help of a curated and labeled dataset and performed with high accuracy of classifying about 95% classes, proving its efficiency in recognizing normal and suspicious activity across static images.

2. RELATED WORK

Advanced theft detection systems apply machine learning methods, like a YOLO model, to reliably and rapidly detect abnormal behavior captured in supervision systems. Besides, these systems are based on the vision of CCTV cameras, for analyzing abnormal behaviors and offering an intrusion detection, which makes them better to monitor and early prevent theft, whether at home, shops, museums, etc. [16]. Computer vision with artificial intelligence and machine learning, as well as the *OpenCV* library, help for highly precise and fast object detection and recognition. Other techniques are detection, recognition and image segmentation, which help to facilitate the processing and analyze images better. Computer vision and ML are integral elements to ensure accurate and fast classification of variety activities like monitoring driver's behavior to enhance safety, reduce the accident that has been affected by environmental reasons (i.e., directions gazing by the driver, traffic on the road and vehicle trajectory) [11]. Deep learning surveillance systems offer intelligent demonstration of object detection and tracking that improve detection performance as well as the quality of the image, in spite of the fact that videos comprise of a vast amount of data. But the work still requires a human to confirm the hits. Things like OpenCV have paved the way for very efficient analysis of live video. In a dynamic system of this type, camera tracking and image analysis are used to detect the position and movement of thieves [17]. The necessity of a monitoring system arises increasing theft, violence problem and traditional systems can have limitations, lack of immediate alert, false alarms at night, difficulties in the identification. The requirement of continuous checking also limits these systems to use by the average user with wise technical assistance [18]. YOLO-FIR, an enhanced variant of YOLO for infrared imagery, was proposed with YOLOv5+ to obtain better detection performance, especially in practical scenes. This model solves the problem of low detection accuracy for small-distant objects and has better detection performance

than the previous YOLO models, which is applicable to real-time video applications [19]. In this related work, they analyzed both YOLOv3 and YOLOv4 for detecting firearm(s) in the real-world context, which is crucial to minimize any potential delay for recognizing these weapons to ensure public safety and security, and to prevent accidents in both public and private areas [20].

3. ETHICAL IMPLICATIONS OF FALSE POSITIVES IN RETAIL SURVEILLANCE

Considering the automated retail surveillance application, there are ethical concerns, specifically regarding biases and implications of false positive detections. One major issue is population bias in the training data – the system could end up unfairly targeting people of a certain age, gender or ethnic profile. This bias can result in discriminatory treatment of shoppers, and it raises critical ethical and legal questions. More generally, false-positive detections are a real concern as they may lead to false incriminations of innocents and may therefore erode public confidence in the surveillance authority. **Methods** We propose a human-in-the-processing loop for flagged cases to minimize this risk, to make model transparent, and to regularly audit and improve the model using diverse and representative datasets. It is crucial to address these ethical issues to ensure fairness, accountability and social acceptance of surveillance systems in the retail industry.

The ethical issues on the proposed system were mitigated by a number of procedures such as: training multiple parameters with diverse dataset to mitigate demographic bias; regular model review with new data to check for drifts in the model behavior; including human feedback early in the decision process; tuning model sensitivity in order to reduce the rate of false positives; keeping track of all outcomes to ensure due process and transparency.

4. PROPOSED SYSTEM

In this paper, we introduce an intelligent system for suspicious behavior classification in the small shopping environment using manually labeled still images through the YOLOv8 algorithm. At the system level, we aim to achieve low computational complexity and low latency to provide a practical surveillance system measuring long-term in-store activity by counting and identifying shoppers and potential shoplifters directly at the frame level, without relying on temporal analysis or behavior modelling.

The stages of proposed system consist of the following:

Data collection: Videos were captured from an experimental environment simulating a real store, and the frames were extracted and converted into still images.

Labelling: A manual learning tool *RoboFlow* was used to add labels to each image, such as (normal, suspicious).

Data preparation for YOLO training: Converting images and labels to the format required by YOLOv8.

Training the YOLOv8 model: Using a dataset containing over (15,000) images for training and over (700) images for evaluation or testing.

Image classification: The trained model detects and classifies objects within the image instantly.

Display results: The classified images are output with the bounding box, label, and classification probability actual

predictions with the ground truth.

In the proposed system suspicious behavior detection system, this research paper focuses on accurate data classification using the YOLOv8 model for object detection within a surveillance environment. The current work involves classifying behavioral activities based on manually labelled images collected from multiple sources, identifying behavioral patterns of individuals within a simulated or real-world environment. The YOLOv8 model is used to analyze images and extract information about individuals and objects in the image. Through accurate object detection, activities are classified into categories such as normal or suspicious, providing preliminary information for understanding the behavior of individuals as they interact with products in stores. Figure 1 illustrates the general structure of the proposed system. The goal of this study is to build a classification system that identifies real-time motor activities and assigns them to specific categories based on image analysis. For example, if behavior such as (hiding the product in the bag) or (placing the product in the pocket) is observed, it is classified as suspicious, helping to improve security within stores. In the future, this work will be expanded to include temporal behavior analysis. In the next research paper, labelled data will be used to build models capable of analyzing the temporal behaviors of individuals over time. This analysis will help track individuals' movement patterns over a specific period of time, enabling the identification of any abnormal changes that may indicate suspicious behavior. The deployed shoplifting detection system that developed based on YOLOv8 was operate on an edge device, namely NVIDIA – Jetson – TX2. At an input resolution of (640 × 640), the system operates at an inference speed of about 25 Frame Per Second. This indicates that the model can make immediate detection and classification in real-time manner on resource - limited devices and is applicable to real - world retail surveillance systems.

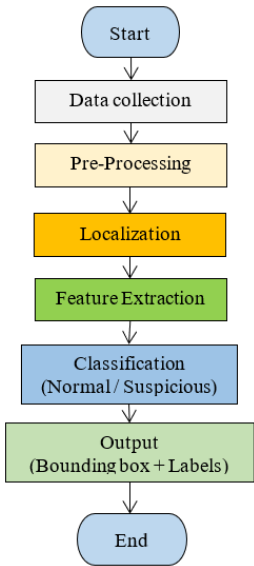


Figure 1. General system architecture

5. EXPERIMENTAL RESULTS

The YOLOv8 model was trained on UCF-Crime data to detect robberies associated with human behavior. This data was augmented with additional images from diverse cases and multiple patterns. The settings used during training were (150)

epochs, a batch size of (16), and an image size of (640 × 640). To ensure the transparency and reproducibility of the proposed methodology system, the training dataset contains (15,919) images manually annotated using the *Roboflow* platform to ensure classification accuracy regarding human-object interactions observed in retail settings. These images are evenly distributed into normal and suspicious categories. Normal behaviors include picking up objects, looking at them, examining them, and returning them to shelves, while suspicious behaviors include picking up objects and placing them in pockets, bags, or clothing. Table 1 shows a balanced image distribution, with each category containing (7,959) images. This even distribution of categories helps reduce bias in the training process, enabling the model to effectively learn the attributes that characterize each behavior.

Table 1. Balancing distribution of training data

Class	Sample Number	Percentage (%)
Normal	7959	50%
Suspicious	7960	50%
Total	15919	100%

We also reported box loss, classification loss, and DFL on the validation set. The test results were similar to the training results, suggesting no overfitting which means no overlearn on the training data. Results from this stage are presented in Figure 2.

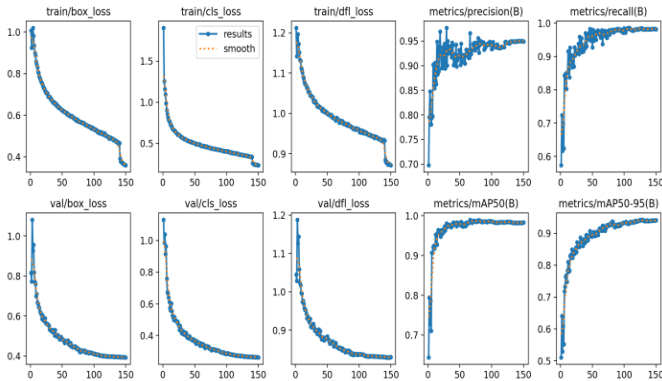


Figure 2. Loss analysis result

5.1 Loss analysis

One of the key metrics for evaluating model performance is loss. In this research, we evaluated three main types of losses:

Train / box loss: This represents the loss resulting from identifying the bounding boxes of objects. We observed a gradual decrease in this value over the epochs, indicating that the model was able to locate objects more accurately with each training cycle.

Train / classification loss: The results suggest the improvement in hits rate (i.e., accuracy) over epochs.

Train / distribution focal loss: We added the DFL to assist the model in more fitting with imbalance distributed object. The loss on the less frequent object-related challenges was substantially smaller, indicating that the model is able to address these as well.

5.2 Precision and recall analysis

Precision: The accuracy of predicting relevant objects was

measured. The results showed a gradual improvement in the model's accuracy with each epoch. Accuracy increased significantly, demonstrating the model's ability to accurately identify objects. Figure 3 illustrates this metric.

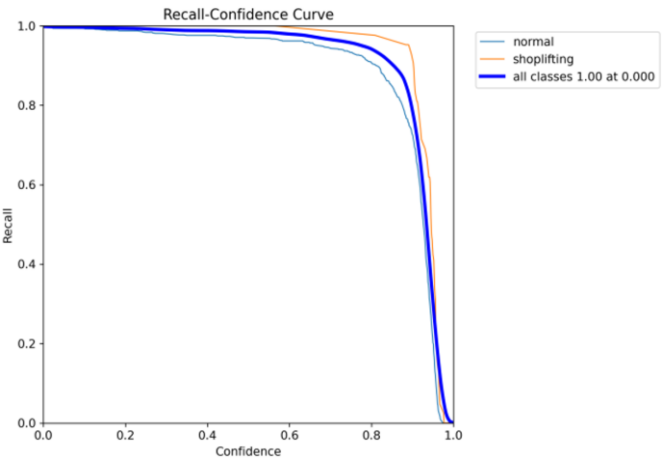


Figure 3. Precision analysis result

Recall: The recall results reflect the model's ability to find the correct objects among all relevant objects. The results showed that recall increased over the epochs, meaning the model was able to detect a greater number of relevant objects. Figure 4 illustrates this metric.

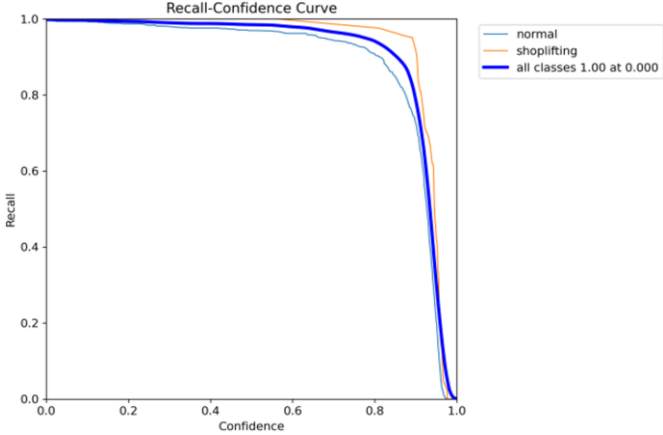


Figure 4. Recall analysis result

mAP50: A measure of the mean average precision at (50%) of the threshold. The results showed that the model was highly efficient at detecting objects, with a high mAP50, reflecting the model's ability to predict objects with good accuracy.

mAP50-95: A more comprehensive measure of model performance across a range of thresholds from (50%) to (95%). These values demonstrated good performance in accurately classifying objects across multiple thresholds. Table 2 shows the results of the last five training epochs. The last epoch showed good results, with an accuracy of nearly (95) percent.

Table 2. The training for last five epochs results

Epoch Number	Metrics/Precision	Metrics/Recall	Metrics/mAP50	Metrics/mAP50-95	Prediction Time (ms)
146	0.94972	0.9829	0.98187	0.93964	13
147	0.94979	0.98285	0.98187	0.9396	15
148	0.94986	0.98227	0.98244	0.93999	17
149	0.94995	0.98216	0.9829	0.93973	14
150	0.94935	0.98196	0.98288	0.94068	12

5.3 Confusion matrix analysis

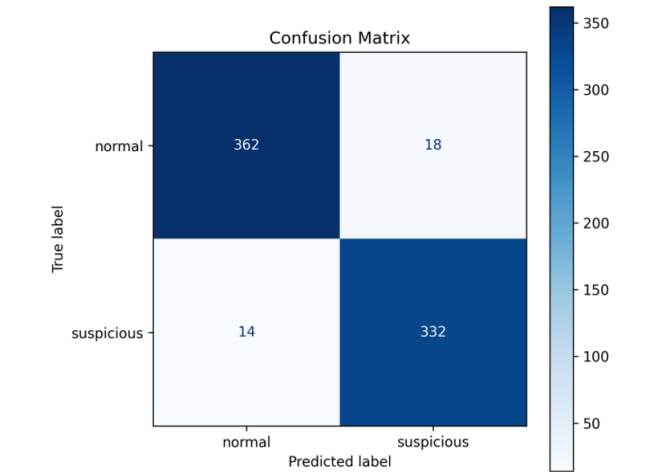


Figure 5. Confusion matrix

To fully validate the results of the proposed theft detection system, not only the confusion matrix but also the classification report were used, as shown in Figure 5 and Table 3. These tools facilitate the deep investigation of the model's predictive performance beyond the aggregate accuracy, in particular, in the case of every class implicated. The approach was validated on a (726) test images belonging to two

categories: normal and suspicious. As can be seen from the confusion matrix, the network could accurately identify 480 of (500) of the normal and (210) of (226) of the suspicious images at approximately (95.03) accuracy. Further validation was conducted with the use of the classification report. These results indicate that the model is robust enough to correctly identify both normal and suspicious behaviors (Figure 6), and it performs especially well in the suspicious class, which is fundamental in real scenarios of theft detection.

Table 3. Report classification

Class	Precision	Recall	F1-score
Normal	0.963	0.953	0.958
Suspicious	0.949	0.96	0.954



(a) Represents the snapshots of the images that were labelled



(b) Represents what the model predicted

Figure 6. Visual results of the system

5.4 Comparison with other models

YOLOv8 was also evaluated against different models in object detection, to wit, YOLOv5, Faster R-CNN and SSD. The numbers indicated that there is a distinct difference among performance of the models, YOLOv8 being the best among the others for most of the metrics below:

YOLOv5: It achieved an mAP@0.5 of about (90%), but (9%) lower recall than YOLOv8.

Faster R-CNN: Although it did achieve relatively good results, it needed more time in prediction, thus, it was not efficient for real-time applications.

SSD: It was the least accurate, but most successful at predicting speed and a small model size.

The final comparison between the models was shown in the Figure 7. These findings testify YOLOv8's better performance in capturing theft-behaviors in realistic scenarios. The method offers high accuracy with execution speed and it is suitable for real time video surveillance application. Furthermore, the large differences in the other models emphasize that the choice of model depends on what is more important (speed / accuracy / size) for the target application. To compare performance consistency across different object detection models, we repeated five independent experimental runs for each model and wrote down their Precision scores in forms of As shown in the Table 4.

Table 6. Paired t-test results between YOLOv8 and other object detection

Comparison	t-value	Degrees of Freedom (df)	p-value	Significance
[YOLOv8, YOLOv5]	12.33	8	<0.001	Significant
[YOLOv8, Faster R-CNN]	22.33	8	<0.001	Significant
[YOLOv8, SSD]	34.00	8	<0.001	Significant

Table 7. YOLOv8 performance across various metrics

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Time (ms)
YOLOv8	0.96	0.97	0.98	0.94	12
YOLOv5	0.89	0.88	0.9	0.85	10
Faster R-CNN	0.83	0.85	0.88	0.81	135
SSD	0.76	0.78	0.82	0.75	15

6. CONCLUSIONS

It is found that YOLOv8 is significantly (6% higher) better

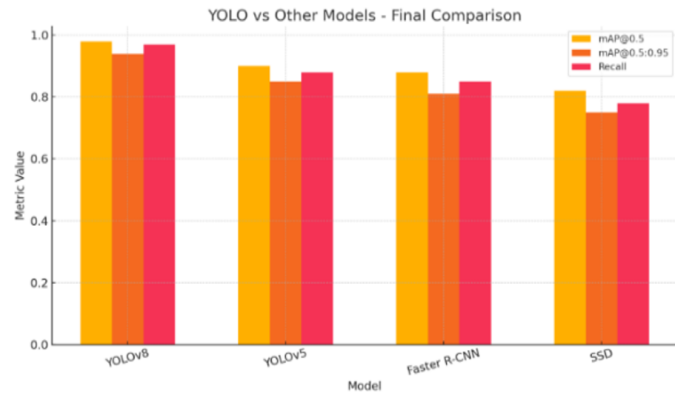


Figure 7. Model comparison

Precision scores between these models were significantly different as confirmed by a one-way (ANOVA) test ($F(3,16) = 342.61$, $p < 0.001$). Specific differences were investigated using a pairwise t-test against the baseline models and YOLOv8. Our findings revealed that YOLOv8 significantly exceed the performances of both YOLOv5 ($t(8) = 12.33$, $p < 0.001$), Faster R-CNN ($t(8) = 22.33$, $p < 0.001$), and SSD ($t(8) = 34.00$, $p < 0.001$) illustrated in Table 5. These results prove the better and statistically significant accuracy of the YOLOv8 model with respect to the other considered detection methodologies as shown in Table 6. Table 7 shows the performance numbers across various metrics of our YOLOv8 model.

Table 4. Precision score values

Model	Pre (1)	Pre (2)	Pre (3)	Pre (4)	Pre (5)
YOLOv8	0.95	0.96	0.97	0.96	0.96
YOLOv5	0.88	0.89	0.90	0.89	0.87
Faster R-CNN	0.82	0.83	0.84	0.83	0.81
SSD	0.75	0.76	0.77	0.76	0.74

Table 5. Mean and standard deviation result

Model	Precision Mean	Std deviation
YOLOv8	0.960	0.00707
YOLOv5	0.886	0.01140
Faster R-CNN	0.826	0.01140
SSD	0.756	0.01140

than other state-of-the-art detection and classification system, especially in criminal behavior for instance theft. The model showed a remarkable change trends during the training

process; it achieved a decrease on loss functions and an increase on performance indicators (Precision, Recall for mAP). Moreover, the confusion matrix and ROC/PR curves showed the model could accurately classify in most regions, even among ambiguous or overlapped classes, which would benefit the reliability of the model in a wide range of practical applications. From a research point of view, these proofs confirm a cornerstone of reasonably intelligent systems supporting real-time detection and understanding of human behaviors, especially in security and surveillance domains. The outputs of this model could be used in future works to combine with temporal modeling models, e.g., Long Short Term Memory (LSTM) or transformer based architectures to track and analyze behavioral evolution across different frames in videos, which would further enhance performance in predicting intent and detecting suspicious behaviors.

REFERENCES

- [1] Smith, T.A. (2020). Investigations: Consumer retail shoplifting. In *Encyclopedia of Security and Emergency Management*, pp. 1-7. https://doi.org/10.1007/978-3-319-69891-5_172-1
- [2] Shrestha, S., Taniguchi, Y., Tanaka, T. (2024). Detection of pre shoplifting suspicious behavior using deep learning. In *2024 16th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, Takamatsu, Japan, pp. 450-455. <https://doi.org/10.1109/IIAI-AAI63651.2024.00088>
- [3] He, K., Gkioxari, G., Dollár, P., Girshick, R. (2017). Mask r-cnn. In *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 2980-2988. <https://doi.org/10.1109/ICCV.2017.322>
- [4] Gim, U.J., Lee, J.J., Kim, J.H., Park, Y.H., Nasridinov, A. (2020). An automatic shoplifting detection from surveillance videos (student abstract). *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(10): 13795-13796. <https://doi.org/10.1609/aaai.v34i10.7169>
- [5] Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., Parmar, M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57(4): 99. <https://doi.org/10.1007/s10462-024-10721-6>
- [6] Zingoni, A., Alcalde-Llargo, J.M., Morciano, G., Melloni, D., et al. (2024). Real-time detection of criminal actions in the everyday life, from camera-equipped streetlamps. In *2024 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRaine)*, St Albans, United Kingdom, pp. 529-534. <https://doi.org/10.1109/MetroXRaine62247.2024.10795879>
- [7] Chen, Y., Yuan, X., Wang, J., Wu, R., Li, X., Hou, Q., Cheng, M.M. (2025). YOLO-MS: Rethinking multi-scale representation learning for real-time object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(6): 4240-4252. <https://doi.org/10.1109/TPAMI.2025.3538473>
- [8] Cai, Y., Luan, T., Gao, H., Wang, H., et al. (2021). YOLOv4-5D: An effective and efficient object detector for autonomous driving. *IEEE Transactions on Instrumentation and Measurement*, 70: 1-13. <https://doi.org/10.1109/TIM.2021.3065438>
- [9] Vaghela, R., Vaishnani, D., Srinivasu, P.N., Popat, Y., Sarda, J., Woźniak, M., Ijaz, M.F. (2025). Land cover classification for identifying the agriculture fields using versions of YOLO v8. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18: 8672-8684. <https://doi.org/10.1109/JSTARS.2025.3547058>
- [10] Patrikar, D.R., Parate, M.R. (2022). Anomaly detection using edge computing in video surveillance system. *International Journal of Multimedia Information Retrieval*, 11(2): 85-110. <https://doi.org/10.1007/s13735-022-00227-8>
- [11] Varun, S., Bhuvanesh, V.M. (2023). Real time theft detection using YOLOv5 object detection model. In *2023 3rd International Conference on Innovative Sustainable Computational Technologies (CISCT)*, Dehradun, India, pp. 1-5. <https://doi.org/10.1109/CISCT57197.2023.10351223>
- [12] Duja, K.U., Khan, I.A., Alsuhaibani, M. (2024). Video surveillance anomaly detection: A review on deep learning benchmarks. *IEEE Access*, 12: 164811-164842. <https://doi.org/10.1109/ACCESS.2024.3491868>
- [13] Talaat, F.M., ZainEldin, H. (2023). An improved fire detection approach based on YOLO-v8 for smart cities. *Neural Computing and Applications*, 35(28): 20939-20954. <https://doi.org/10.1007/s00521-023-08809-1>
- [14] Al-Iqubaydhi, N., Alenezi, A., Alanazi, T., Senyor, A., et al. (2024). Deep learning for unmanned aerial vehicles detection: A review. *Computer Science Review*, 51: 100614. <https://doi.org/10.1016/j.cosrev.2023.100614>
- [15] Chopuk, P., Boonmee, P., Jiarasuwan, S., Jearasuwan, S., Bookprakong, P. (2023). Theft detection by patterns of walking behavior using motion-based artificial intelligence. In *International Workshop on Advanced Imaging Technology (IWAIT)*, Jeju, Korea, pp. 140-145. <https://doi.org/10.1117/12.2671245>
- [16] Sharma, A., Pathak, J., Prakash, M., Singh, J.N. (2021). Object detection using OpenCV and python. In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, Greater Noida, India, pp. 501-505. <https://doi.org/10.1109/ICAC3N53548.2021.9725638>
- [17] Khadse, S., Nandanwar, B., Kirme, A., Bagde, T., Kamble, V. (2022). Machine learning based theft detection using yolo object detection. *International Journal of Scientific Research in Science and Technology*, 9(1): 117-120.
- [18] Pandya, S., Ghayvat, H., Kotecha, K., Awais, M., et al. (2018). Smart home anti-theft system: A novel approach for near real-time monitoring and smart home security for wellness protocol. *Applied System Innovation*, 1(4): 42. <https://doi.org/10.3390/asi1040042>
- [19] Li, S., Li, Y., Li, Y., Li, M., Xu, X. (2021). YOLO-FIRI: Improved YOLOv5 for infrared image object detection. *IEEE Access*, 9: 141861-141875. <https://doi.org/10.1109/ACCESS.2021.3120870>
- [20] Hashmi, T.S.S., Haq, N.U., Fraz, M.M., Shahzad, M. (2021). Application of deep learning for weapons detection in surveillance videos. In *2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2)*, Islamabad, Pakistan, pp. 1-6. <https://doi.org/10.1109/ICoDT252288.2021.9441523>