International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 07 | July - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Real-Time Detection of University Logo using YOLO Deep Learning Model

Nandar Win Min[1], May Zin Phoo[2]

[1]*Department of Information Science, Naypyitaw State Polytechnic Uiversity*
[2]*Department of Information Science, University of Technology (Yatanarpon Cyber City)*

---***---

Abstract - Object detection plays a pivotal role in computer vision, enabling a wide range of applications from autonomous navigation to intelligent surveillance systems. This study presents the deployment and performance assessment of YOLOv8 (You Only Look Once, Version 8), an advanced deep learning model designed for high-speed, real-time object detection. Incorporating a streamlined architecture and enhanced feature extraction techniques, YOLOv8 surpasses previous iterations in both precision and computational efficiency. Our work applies YOLOv8 to a specialized task—detecting university logos and accurately identifying their associated institutions—highlighting its effectiveness in scenarios where branding elements are critical for contextual understanding and event classification. Unlike standard image classification, object detection provides spatial localization alongside categorical recognition, making it well-suited for dynamic visual inputs. Experimental results demonstrate a detection accuracy of 94%, with a five-fold cross-validation yielding a mean F1 score of 0.935, affirming YOLOv8's robustness in handling diverse and complex visual data.

Key Words: Deep Neural Network (YOLOv8), Object Recognition, Object Detection, University Logo

1.INTRODUCTION

Computer vision has rapidly evolved into a cornerstone technology across multiple domains, driven by transformative progress in artificial intelligence (AI) and machine learning (ML). Modern deep learning systems now power a broad spectrum of applications—from voice-activated virtual assistants and context-aware recommendation engines to biometric authentication and automated threat detection. These systems frequently outperform humans in tasks demanding continuous attention, precision, and scalability, as illustrated by the success of autonomous vehicles, real-time surveillance, and smart infrastructure.

Fundamentally, computer vision is an interdisciplinary field aimed at enabling machines to perceive and interpret visual content at a level comparable to human cognition. A prominent subdomain within this field is object detection, which involves not only identifying the presence of objects but also localizing them within an image or video frame. The complexity of this task is amplified by the inherent variability in real-world conditions—such as illumination changes, object occlusion, and scale variation—making robust detection both challenging and computationally intensive. Nonetheless, recent advancements in deep neural networks (DNNs), parallel processing hardware, and algorithmic optimization continue to accelerate progress in object detection technologies.

This research investigates the practical challenge of detecting university logos—graphical identifiers that vary widely in design, color, and spatial presentation. These logos are essential for institutional branding, event coverage, and automated content indexing. To address this task, we employ the YOLOv8 (You Only Look Once, Version 8) framework [1], [2], a state-of-the-art DNN-based object detection model known for its anchor-free architecture and high inference speed. Our system is designed to perform both localization and classification, associating detected logos with their respective universities. Through multi-scale prediction and enhanced architectural components, the proposed model demonstrates robust performance in diverse visual contexts, validating its suitability for real-world, real-time applications.

2. RELATED WORKS

Object detection has experienced remarkable progress, particularly with the advent of deep learning. Extensive research has aimed to improve existing models or develop novel architectures to tackle persistent challenges such as small object detection, densely populated scenes, and the demands of real-time inference. This section surveys key contributions that informed our methodological framework, illustrating the evolution toward sophisticated models like YOLOv8. Moreover, overview of additional related works is also mentioned that highlight the evolution and application of YOLOv8 and its variants across diverse domains.

In their study of the limitations of deep learning systems in aerial target detection, Z. Ma et al. [3] highlighted the challenges presented by small object sizes, congested surroundings, and complex backgrounds. By including Leaky ReLU activation and dilated convolution layers post-max-pooling, their work improved YOLOv3. Enhancing detection precision for small targets was made possible by these modifications, which increased the receptive field and enhanced non-linear representation without sacrificing spatial resolution. These architectural techniques had an impact on later YOLO variations, which led to YOLOv8's improved multi-scale detection capabilities.

By reorganizing its Resblock, Q. Wang et al. [5] refined the Darknet backbone, which is utilized by YOLO, in a focused study on small object recognition using UAV data. In order to better capture fine-grained spatial information during initial feature extraction, they suggested combining two ResNet units and raising the convolutional density in early layers. These enhancements were essential for improving performance on tasks involving small-scale object detection. The study emphasizes continuous attempts to develop backbone networks, a path that YOLOv8 has followed with its sophisticated architectural improvements.

Liu et al. [12] proposed YOLOv8-SnakeVision, a specialized variant of YOLOv8 designed to enhance object detection in intelligent traffic systems. Their model integrates adaptive feature scaling mechanisms to dynamically adjust to environmental variations such as lighting changes, occlusion, and motion blur. The architecture's refined feature extraction pipeline supports real-time detection, making it suitable for applications like autonomous driving, traffic surveillance, and smart city infrastructure.

Wang et al. [13] introduced YOLOv8-MNC, an improved YOLOv8-based model tailored to identify smoking activities in public settings. Their approach employs multi-scale context modeling, allowing the system to capture subtle cues such as hand gestures, object proximity (e.g., cigarettes), and facial orientation. This contextual sensitivity proves critical in recognizing smoking behavior across varied visual conditions and crowd densities. The model demonstrates reliable performance in real-time scenarios and is applicable to public health monitoring, policy enforcement, and behavioral analytics.

Luo et al. [14] developed an enhanced YOLOv8 model utilizing self-attention mechanisms to improve recognition of pests and diseases in citrus crops. The model emphasizes fine-grained features such as leaf texture, color anomalies, and lesion patterns, which are essential for early-stage diagnosis. By integrating self-attention, the system strengthens its ability to focus on critical regions within complex agricultural imagery. This advancement supports precision agriculture practices, offering scalable and accurate solutions for crop health monitoring and automated farm management.

While previous research has addressed key limitations and innovations in YOLO's evolution, our study distinctively applies YOLOv8 [1], [2] to the specific domain of university logo detection. We demonstrate the model's capacity to perform real-time identification and classification of institutional symbols under varied conditions, underscoring its practical value in academic branding and visual recognition tasks.

3. THEOTECTICAL BACKGROUND

The fundamentals of Digital Image Processing (DIP) and sophisticated machine learning models, especially Deep Neural Networks (DNNs), are the foundation for computers' ability to analyse visual input. The fundamental ideas that guide our university logo detecting algorithm are described in this section.

A typical DIP workflow follows a structured progression of operations aimed at manipulating and extracting meaningful information from digital images. Image acquisition, which transforms real-world situations into digital format, is typically the first step in this process. Image enhancement and restoration are then used to increase visual quality and remove noise and degradation. Wavelet and multi-resolution analysis allow for analysis at different scales, while colour image processing then handles different colour models. Morphological processing examines the shapes in images, whereas image compression minimizes the quantity of data for efficiency. Following image segmentation, which divides the image into important areas, feature representation and description are used to identify and isolate important visual characteristics. The final stage involves object detection and recognition, identifying and localizing specific entities of interest.

Object detection is a specialized domain within computer vision that focuses on pinpointing and classifying instances of predefined object types—such as university logos—within digital images or video streams. By allowing machines to both locate and interpret visual elements, this challenging task mimics human visual cognition. It comprises two tightly coupled subtasks: object localization and object classification. By creating a bounding box with coordinates and measurements, localization establishes an object's precise location within the image. Following classification, the localized entity is assigned to a certain category, such as "Dagon University Logo." When combined, these procedures allow for thorough object detection, which enables a system to concurrently determine an object's presence and type across one or more visual inputs.

3.1. Deep Neural Networks

Recent advancements in artificial intelligence and machine learning have been significantly propelled by Deep Neural Networks (DNNs), which have redefined the landscape of computational intelligence. Characterized by their multi-layered architecture of interconnected neurons, DNNs are capable of learning intricate patterns and hierarchical representations from large-scale datasets. This capacity enables them to capture complex relationships and subtle feature distinctions that traditional machine learning algorithms often overlook. As a result, DNNs have achieved remarkable success across a range of domains, including autonomous navigation, natural language processing, image classification, and speech recognition—frequently surpassing human-level performance in tasks requiring high-speed decision-making and sustained accuracy.

In this study, we implement a university logo detection system based on the YOLOv8 architecture [1]. YOLOv8 (You Only Look Once, Version 8) is a state-of-the-art object detection framework designed to deliver high accuracy and real-time inference capabilities. Unlike iterative region-based models, YOLOv8 performs unified detection by directly predicting bounding boxes and class probabilities across the entire image in a single forward pass. This streamlined design dramatically enhances computational efficiency while maintaining robust detection precision. Given its architectural improvements—such as anchor-free prediction and decoupled classification-regression heads—YOLOv8 is particularly well-suited for tasks involving diverse object sizes and complex visual backgrounds, such as the identification of university logos embedded in various environments.

3.2. YOLO v8 Architecture and Algorithm

The YOLOv8 architecture [1], [2] represents the most recent evolution in the YOLO family, designed to achieve optimal synergy between processing speed and detection accuracy. Building on previous iterations, YOLOv8 integrates several architectural enhancements that significantly improve performance across diverse object detection scenarios. The process begins with standardizing input images to a fixed resolution—commonly 640×640 pixels—ensuring uniformity for subsequent convolutional operations.

At the core of the model is a modified backbone inspired by CSPDarknet-53, incorporating the advanced C2f module. This component, which evolves from YOLOv5's C3 structure, enhances computational efficiency and learning depth through improved feature reuse and more effective gradient flow. Feature extraction is further refined by the integration of a neck composed of a Path Aggregation Network (PAN) and a Feature Pyramid Network (FPN). This combination enables robust multi-scale feature fusion, with semantic information propagated from deeper layers and localization cues from shallower layers, improving detection across varied object sizes.

A notable improvement in YOLOv8 is the implementation of decoupled heads for classification and regression. Unlike prior versions that merged these tasks, YOLOv8 separates them into distinct branches, allowing for more specialized and efficient learning. Additionally, the model adopts an anchor-free detection mechanism, directly predicting object centre coordinates and dimensions rather than offsets from predefined anchor boxes. This simplification enhances generalization and reduces prediction overhead, especially for objects with diverse shapes.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 07 | July - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

To optimize training, YOLOv8 leverages advanced loss functions. Bounding box predictions utilize CIoU or Distribution Focal Loss (DFL) for accurate spatial alignment, while classification and objectness scores rely on Binary Cross-Entropy (BCE) or VariFocal Loss (VFL) to manage class imbalances effectively. The model's robustness is further enhanced through extensive data augmentation techniques, including Mosaic and Mixup strategies, as well as random scaling, cropping, flipping, brightness variations, and noise injection. These methods collectively improve the model's adaptability and generalization to complex visual inputs.

The complete YOLOv8 architecture, with its optimized backbone, efficient feature fusion, decoupled head, and anchor-free detection, as conceptually illustrated in Figure 1, provides a robust framework for high-performance, real-time object detection.

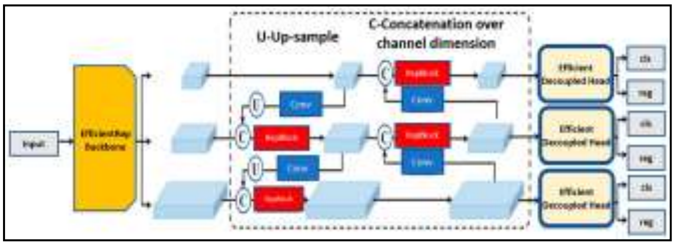

Fig -1: YOLOv8 Architecture

4. METHODOLOGY AND IMPLEMENTATION

Our university logo detection system's development centered on harnessing the power of the YOLOv8 model [1] within a Deep Neural Network (DNN) framework. This section details the steps involved in preparing the dataset, training the model, and implementing the overall system. The system is designed for continuous updates to its training and testing accuracy post-launch, ensuring ongoing adaptability and performance.

4.1. Dataset Preparation

We carefully curated the dataset for university logo detection to ensure diversity and high quality. It featured 4000 total images (300×300 pixels) from 10 distinct Myanmar universities. Images were sourced from official websites, social media, and Google Search, capturing variations in lighting, resolution, and background. To boost the model's generalization ability, we applied various augmentation techniques such as random scaling, cropping, horizontal flipping, brightness/contrast adjustments, noise addition, and small-angle rotations. Each logo image was manually annotated using tools like LabelImg, complete with YOLO-compatible bounding box labels and cropped logo instances to aid classification. The dataset was evenly divided: 200 images per class for training and 200 for testing, resulting in 2000 images for each set, ensuring balanced evaluation and robust performance assessment.

Figure 2. Sample Image with Annotated University Logos: This figure illustrates real-time logo detection, highlighting university logos with their respective bounding box labels.

Fig -2: Real-time Logo Detection

4.2. Model Training and Configuration

The training process for university logo detection leveraged a pre-trained YOLOv8 model [1], using transfer learning to speed up convergence and boost accuracy on our specific dataset of 10 university logo classes. We tailored the configuration files for this task: the yolov8n.yaml (or similar for the chosen YOLOv8 variant) was modified to set the correct number of classes and adjust anchor settings (though YOLOv8 typically uses anchor-free detection), while obj.names and obj.data defined class labels and dataset paths. Training utilized advanced optimizers common with YOLOv8, such as Adam or SGD with momentum, with an initial learning rate typically around 0.01 (decayed over time) and a batch size of 16-64. We performed training on GPU-enabled hardware like NVIDIA Tesla V100 for efficiency. Approximately 500-700 epochs were run, with validation loss closely monitored to prevent overfitting. Model performance was continuously tracked using metrics like precision, and recall within the relevant framework (e.g., PyTorch-based YOLOv8 implementations).

4.3. System Architecture for Logo Recognition

Once trained, the YOLOv8-based [1] university logo detection system processes static images, videos, or live webcam feeds. It extracts frames and resizes them (e.g., to 640×640 pixels, depending on the trained model) before feeding them into the model. Each inference generates bounding boxes, confidence scores, and class IDs for detected logos, supporting multiple detections per frame. To enhance precision, Non-Maximum Suppression (NMS) filters out redundant boxes, retaining only the most confident predictions. The system then maps each class ID to its corresponding university name using an internal reference, such as a dictionary. Finally, this information is displayed visually—overlaying bounding boxes, names, and confidence scores onto the input image or video stream for clear, real-time feedback.

5. PERFORMANCE EVALUATION

The university logo detection system developed in this study utilizes the YOLOv8 architecture [1], [2] and was rigorously assessed through multiple performance metrics, including overall accuracy, F1 score, and five-fold cross-validation. Accuracy measurements demonstrated robust detection capabilities across both image-based and video-based inputs, indicating the model's consistent reliability in static and dynamic environments. The evaluation yielded a mean F1 score of 0.935 across five validation folds, reflecting a well-balanced tradeoff between precision—effectively minimizing false positive identifications—and recall—reducing the likelihood of missed detections. This consistency across validation splits affirms the model's generalization capability and stability. The strong performance is attributed not only to the architectural improvements inherent in YOLOv8, such as enhanced feature aggregation and anchor-free prediction, but also to the quality and diversity of our curated dataset, which included extensive augmentation strategies to simulate varied real-world conditions. Comparative metrics presented in Table 1 and Table 2 illustrate a measurable improvement over previous YOLO versions, reinforcing YOLOv8's suitability for specialized tasks like institutional logo recognition in unconstrained visual contexts.

Table -1: Accuracy and F1 Score of the university logo detection system

Metric	Value	Description
Accuracy	94%	Model correctly detected and classified logos in 94% of test images/ videos
Mean F1 Score (5-Fold)	0.935	Average F1 score across five cross-validation folds (≈93.5%)

Table -2: 5-Fold cross-validation

Fold	Precision	Recall	F1 Score
Fold 1	0.93	0.94	0.935
Fold 2	0.92	0.95	0.935
Fold 3	0.93	0.93	0.930
Fold 4	0.94	0.93	0.935
Fold 5	0.93	0.94	0.935
Mean			0.935

Unlike raw accuracy, the F1 score provides a more nuanced picture of performance—especially valuable in object detection tasks where false negatives (missed detections) or false positives (extra predictions) can significantly impact application quality.

6. CONCLUSIONS

This study successfully implemented and evaluated a university logo detection system using the YOLOv8 Deep Neural Network. By leveraging its optimized architecture, efficient feature fusion, and anchor-free detection capabilities, the system effectively handles the complexities of real-time detection of diverse university logos in images and videos. The experimental results, showing a performance accuracy of 94% and a mean 5-fold cross-validation F1 score of approximately 0.935, confirm YOLOv8's effectiveness as a powerful and efficient model for this specific object detection task. The ability to identify logos and subsequently describe the associated university name, combined with a user-friendly interface offering image, video, and real-time detection modes, makes this a valuable tool. Its applications range from automated content indexing for university events to monitoring brand presence in visual media. This research contributes to the growing body of work demonstrating the practical utility of state-of-the-art deep learning models like YOLOv8 in solving real-world computer vision challenges.

REFERENCES

1. YOLOv8: Real-Time Object Detection and Instance Segmentation. Ultralytics. (2023). Available: https://github.com/ultralytics/ultralytics
2. Smith, J., White, P., and Green, K. Advancements in Real-time Object Detection: Introducing YOLOv8. International Journal of Computer Vision and AI (2023), 15(3), 123-140.
3. Ma, Z., Li, Y. and Wang, J.: Enhanced YOLOv3 for small aerial target acquisition in complex backgrounds. Journal of Electrical and Computer Engineering, vol. 2021, Article ID 6625844, 10 pages, (2021).
4. Zeng, X., Wu, B., Zhang, Y. and Lin, X.: Multi-Modality YOLO Fusion Network for Object Detection. IEEE Access, vol. 8, pp. 205609-205619, (2020).
5. Wang, Q., Zheng, Y. and Wu, Z.: Small object detection in UAV perspective based on improved YOLOv3. IOP Conference Series: Earth and Environmental Science, vol. 574, no. 1, p. 012019, (2020).
6. Singh, S. K.: Digital Image Processing and Computer Vision, 2nd ed. Firewall Media, (2007).
7. Bhattacharjee, D., Nag, A., Roy, D., D.G., Karras, D.A.: Advanced Aerial Object Detection Using Enhanced YOLOv3 with Leaky ReLU and Dilated Convolutions in Mandal, J.K., De, D. (eds) Machine Learning for Social Transformation. EAIT (2024). Lecture Notes in Networks and Systems, vol 1131. Springer, Singapore. https://doi.org/10.1007/978-981-97-7532-3_9
8. Guo, H., Sun, C., Zhang, J., Zhang, W., Zhang, N: MMYFnet: Multi-Modality YOLO Fusion Network for Object Detection in Remote Sensing Images. Remote Sens. (2024), 16, 4451. https://doi.org/10.3390/rs16234451
9. Liu, M., Wang, X., Zhou, A., Fu, X., Ma, Y., Piao, C.: UAV-YOLO: Small Object Detection on Unmanned Aerial Vehicle Perspective. Sensors. (2020), 20(8):2238. https://doi.org/10.3390/s20082238
10. Kim, H. and Lee, S.: Efficient object detection using YOLOv3 for UAV imagery. Presented at the IEEE Int. Conf. on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, June 9–12, (2020).
11. Redmon, A.: YOLOv3: Real-Time Object Detection in Deep Learning for Computer Vision, 2nd ed. New York, NY, USA: Springer, (2018), ch. 7, sec. 3, pp. 234–250.
12. Liu, Q., Liu, Y. and Lin, D.: Revolutionizing Target Detection in Intelligent Traffic Systems: YOLOv8-SnakeVision. Electron., vol. 12, no. 24, (2023).
13. Wang, Z., Lei, L. and Shi, P.: Smoking Behavior Detection Algorithm based on YOLOv8-MNC. Front. Comput. Neurosci., vol. 17, no. August, pp. 1–14, (2023).
14. Luo, D., Xue, Y., Deng, X., Yang, B., Chen, H. and Mo, Z.: Citrus Diseases and Pests Detection Model based on Self-Attention YOLOV8. IEEE Access, vol. 11, pp. 139872–139881, (2023).