2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
ASEE AMERICAN SOCIETY FOR ENGINEERING EDUCATION
Paper ID #53517

# Toward a Unified Campus Complexity Index for Object Detection in Surveillance Environments: A Framework Synthesis and Proof-of-Concept Validation

**Mr. Fabrice Simpore, The University of Oklahoma**

Fabrice Simpore is an undergraduate mechanical engineering student at the University of Oklahoma (BS expected May 2026). He plans to pursue graduate studies in mechanical engineering at OU. Contact: fabrices26@ou.edu

**Dr. Moses Olayemi, The University of Oklahoma**

Dr. Mo (Moses Olayemi) is an Assistant Professor of Engineering Pathways at the University of Oklahoma. He is the Founding President of the African Engineering Education Fellows in the Diaspora, a non-governmental organization that leverages the experiences of African scholars in engineering education to inform and support engineering education policy, practice, and pedagogies in Africa. His research revolves around culturally relevant engineering education, STEM education in resource-limited contexts, instrument development, study abroad outcomes and processes, mixed methods research, course design and alignment of content, assessment, and pedagogy, comparative analyses of engineering education systems, and the food-energy-water nexus.

# Toward a Unified Campus Complexity Index for Object Detection in Surveillance Environments: A Framework Synthesis and Proof-of-Concept Validation


## Abstract

Object detection systems deployed in campus surveillance environments face documented failure modes. Crowd density causes occlusion. Poor camera quality degrades features. Variable lighting obscures objects. Motion blur distorts boundaries. Temporal anomalies shift baseline expectations. The computer vision literature identifies each factor independently, yet no unified framework exists to quantify their combined effect on detection difficulty. This paper synthesizes findings from peer-reviewed studies to propose the Campus Complexity Index (CCI), a five-component framework combining established failure modes into a single computable metric. We describe the literature selection methodology, the coding process for extracting failure modes, the mathematical formulation of each component, and the rationale for their combination. To demonstrate feasibility, we validate on two public datasets: UCF CC 50 (50 crowd images) and the Monash Guns Dataset (7,780 weapon detection images), showing that CCI produces meaningful variation between scenario types. Full empirical validation comparing detection performance with CCI-adaptive versus static thresholds remains future work.



**Keywords:** object detection, campus surveillance, adaptive thresholds, literature synthesis, framework development


## Introduction

Universities deploy surveillance cameras across buildings, parking structures, and outdoor spaces. These systems must operate under varying conditions: crowded walkways between classes, dim lighting at night, aging camera hardware, and activity patterns that shift throughout the day. A detector tuned for one condition may fail in another. This paper proposes a framework for quantifying scene difficulty to support detector evaluation and threshold tuning.

Deep learning has advanced object detection significantly. Modern architectures achieve high accuracy on benchmarks like MS COCO and PASCAL VOC. However, these benchmarks use curated images with controlled conditions. When detectors are deployed in real surveillance environments, performance drops. The drop follows predictable patterns: detectors struggle with crowded scenes where people overlap, fail on blurry or noisy camera feeds, miss objects in low light, and lose targets during fast motion.

Each failure mode is documented in the literature. Wang et al. (2018) found that 48.8% of pedestrians in urban scenes overlap with another person, and crowd occlusion accounts for 60% of missed detections. Hendrycks and Dietterich (2019) showed that detectors

lose significant accuracy under common corruptions like blur, noise, and compression. Zhu et al. (2017) reported detection accuracy dropping from 82% on slow-moving objects to 51% on fast-moving ones. Loh and Chan (2019) found that fewer than 2% of training images are captured in low-light conditions. Sultani et al. (2018) showed that anomalous events correlate with detection errors.

The problem is that these findings remain scattered. A practitioner evaluating a detector for campus surveillance has no unified way to characterize scene difficulty. A researcher designing complexity-aware training has no standardized metric. An engineer tuning thresholds has no principled basis for adaptation. This paper addresses that gap by synthesizing findings from peer-reviewed studies into a five-component Campus Complexity Index. Each component corresponds to an established failure mode: crowd density, camera quality, motion, lighting, and temporal patterns. The index can be computed from video frames using standard image processing.

Our contribution is twofold. First, we provide a structured synthesis with clear mathematical formulation. Second, we demonstrate feasibility by computing CCI on two real-world datasets, showing the framework produces varying complexity scores across scenario types. The methodology follows established practices for scoping reviews in engineering education research (Chen, 2025; Borrego et al., 2014).

## Literature Review and Methodology

We followed a structured approach modeled on scoping reviews in engineering education (Chen, 2025; Borrego et al., 2014). We searched IEEE Xplore, ACM Digital Library, arXiv, Scopus, and Google Scholar using terms combining object detection with failure, degradation, robustness, occlusion, crowd, low light, motion blur, and anomaly detection. We limited the search to papers published between 2016 and 2026 to capture the deep learning era.

Papers were included if they reported quantitative detection performance under at least one failure condition, were peer-reviewed or widely-cited preprints, and focused on visual object detection. Papers were excluded if they addressed only classification, focused exclusively on autonomous driving, or reported only qualitative observations. After screening, the final set provided evidence for each of the five failure modes.

We developed a codebook defining what evidence qualifies for each failure mode category, following content analysis practices from Wilhelmsen and Dixon (2016). Table 1 presents the codebook definitions.

*Table 1: Codebook Definitions for Failure Mode Categorization*

| Code | Definition | Qualifying Evidence |
|---|---|---|
| CDI | Detection failure from crowding or occlusion | Miss rate or mAP comparing density conditions |
| CQI | Detection failure from image corruption or blur | mAP drop under controlled corruption |
| MRI | Detection failure from object motion or blur | mAP stratified by motion speed |
| LVI | Detection failure from low-light or illumination | Performance comparison day vs. night |
| TODI | Detection based on deviation from normal activity | AUC comparing normal vs. anomalous |

Each paper was read in full. Sections reporting experimental results and failure analysis were highlighted. Quantitative claims were extracted and matched against codebook definitions. Table 2 presents representative excerpts justifying each coding assignment.

*Table 2: Representative Coding Evidence by Failure Mode Category*

| Code | Source | Key Finding | Metric |
|---|---|---|---|
| CDI | Wang 2018 | 48.8% pedestrians overlap; 60% misses from crowds | Miss rate |
| CDI | Pang 2019 | Miss rate 44.2% to 39.4% on heavily occluded set | Miss rate |
| CQI | Hendrycks 2019 | Significant mAP drop under 75 corruption types | mCE |
| MRI | Zhu 2017 | mAP 82.4% (slow) to 51.4% (fast motion) | mAP |
| LVI | Loh 2019 | <2% of training images are low-light | Training % |
| LVI | Cui 2021 | 11 mAP point gain on ExDark with MAET | mAP |
| TODI | Sultani 2018 | 75.41% AUC on UCF-Crime vs 50% baseline | AUC |
| TODI | Tian 2021 | Magnitude 53.4 (abnormal) vs 7.7 (normal) | Magnitude |

The findings converge on a consistent pattern. Detection degrades along five dimensions with substantial magnitude: 60% of misses from crowd occlusion, mAP dropping from 82% to 51% with fast motion, significant losses under image corruption, near-total failure in extreme low light, and order-of-magnitude differences in anomaly indicators. Several papers proposed adaptive mechanisms, including density-aware NMS (Liu et al., 2019) and motion-stratified evaluation (Zhu et al., 2017), indicating the field is moving toward complexity-aware detection but lacks a unified framework.

## The Campus Complexity Index Framework

Based on the extracted findings, we propose a five-component framework for quantifying scene difficulty. Three considerations guide the design: each component must be computable from video frames using standard image processing, each must correspond to a distinct measurable property, and each must be justified by quantitative evidence from peer-reviewed studies.

The five components, each normalized to [0, 1] where higher values indicate greater complexity, are: Crowd Density Index (CDI), which quantifies crowding and occlusion, justified by Wang et al. (2018) and Pang et al. (2019). Camera Quality Index (CQI) quantifies blur, noise, and compression, justified by Hendrycks and Dietterich (2019). Motion Randomness Index (MRI) quantifies movement unpredictability, justified by Zhu et al. (2017). Luminance Variability Index (LVI) quantifies lighting instability, justified by Loh and Chan (2019). Temporal Outlier Deviation Index (TODI) quantifies deviation from baseline activity, justified by Sultani et al. (2018) and Tian et al. (2021).

The components combine into a single index: CCI = w1*CDI + w2*CQI + w3*MRI + w4*LVI + w5*TODI. Weights sum to one; default weights are uniform at 0.2 each but can be adjusted for specific deployments. We use linear combination for interpretability; alternative aggregation methods are candidates for future comparison.

## Proof-of-Concept Validation

To demonstrate feasibility, we computed CCI on two public surveillance datasets. UCF CC 50 (Idrees et al., 2013) contains 50 crowd images with counts ranging from 96 to 4,633 persons, representing varying crowd densities. The Monash Guns Dataset (Lim et al., 2021) contains 7,780 weapon detection images with bounding box annotations. We sampled 200 images from Monash for efficiency.

CQI was computed using Laplacian variance for blur and local variance for noise. LVI was computed from luminance statistics, including standard deviation and contrast ratio. For static images, MRI was estimated from optical flow between consecutive batch images, and TODI was set to baseline. All computations used OpenCV. Table 3 presents results.

*Table 3: CCI Component Values on Proof-of-Concept Datasets*

| Dataset | N | Avg CQI | Avg LVI | Avg CCI |
|---|---|---|---|---|
| UCF CC 50 (Crowd) | 50 | 0.27 | 0.63 | 0.493 |
| Monash Guns (Weapon) | 200 | 0.23 | 0.62 | 0.385 |
| Difference | - | - | - | +0.108 |

Crowd scenes yielded an average CCI of 0.493 (N=50), while weapon detection scenarios yielded 0.385 (N=200). The difference of 0.108 reflects complexity differences between dense crowd scenes with overlapping subjects and typical weapon scenarios with fewer subjects per frame.

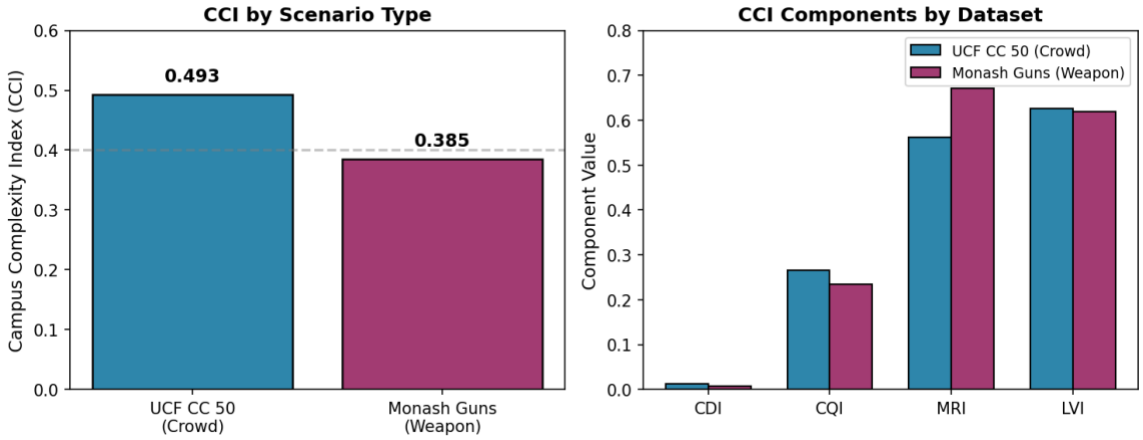


*Figure 1: (Left) CCI by scenario type. (Right) CCI components by dataset.*

As shown in Figure 1, CCI differs between scenario types. The higher CCI for crowd scenes aligns with literature findings that crowd density drives detection difficulty. This validates computational feasibility; full comparison of detection performance with CCI-adaptive versus static thresholds remains future work.

## Discussion and Future Work

This paper synthesizes scattered findings about detection failure modes into a unified framework with clear definitions, computable components, and a combined index. The proof-of-concept demonstrates that CCI can be computed on real imagery and produces values that differentiate scenario types. The framework does not claim to improve detection directly; it provides a structure for reasoning about scene complexity that could support evaluation, adaptation, and training.

The approach may be useful to other researchers facing similar problems. We started with a gap: papers documenting detection failures separately but no way to combine them. We built a codebook, coded the literature, and synthesized results into a unified index. This follows the process Chen (2025) describes for literature methods and Borrego et al. (2014) outline for systematic reviews. Students in other fragmented domains can apply the same steps: identify scattered findings, develop coding criteria, extract evidence systematically, and build a framework from the pieces.

The framework has limitations. Full validation comparing detection performance with CCI-adaptive versus static thresholds remains incomplete. Component computations are simplified. Weights are not optimized. Motion and temporal components are approximated for static images. Coding was performed by one researcher without inter-rater reliability. Future work should run detectors on these datasets with both threshold approaches and compare false positive rates and recall.

## Conclusion

Object detection fails in predictable ways. The literature documents each failure mode, but no unified framework combines them. This paper proposes the Campus Complexity Index, a five-component framework synthesized from peer-reviewed studies. We demonstrate feasibility by computing CCI on 250 real images from two public datasets, finding crowd scenes score higher (CCI = 0.493) than weapon detection scenarios (CCI = 0.385). Full comparison of detection performance remains future work. The contribution is methodological: translating scattered findings into a unified, computable structure.

## Acknowledgments

This work was supported by the Tomorrow's Engineers Research Fellowship at the University of Oklahoma.

## References

Borrego, M., Foster, M. J., & Froyd, J. E. (2014). Systematic literature reviews in engineering education. Journal of Engineering Education, 103(1), 45-76.

Chen, Q. (2025). Demystifying the selection of a literature method. ASEE Annual Conference.

Cui, Z., et al. (2021). Multitask AET with orthogonal tangent regularity for dark object detection. ICCV, 2553-2562.

Hendrycks, D., & Dietterich, T. (2019). Benchmarking neural network robustness to common corruptions. ICLR.

Idrees, H., Saleemi, I., Seibert, C., & Shah, M. (2013). Multi-source multi-scale counting in extremely dense crowd images. CVPR, 2547-2554.

Lim, J., et al. (2021). Deep multi-level feature pyramids: Application for non-canonical firearm detection. Engineering Applications of AI, 97, 104094.

Liu, S., Huang, D., & Wang, Y. (2019). Adaptive NMS: Refining pedestrian detection in a crowd. CVPR, 6459-6468.

Loh, Y. P., & Chan, C. S. (2019). Getting to know low-light images with the ExDark dataset. CVIU, 178, 30-42.

Michaelis, C., et al. (2019). Benchmarking robustness in object detection. NeurIPS Workshop.

Pang, Y., et al. (2019). Mask-guided attention network for occluded pedestrian detection. ICCV, 4967-4975.

Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. CVPR, 6479-6488.

Tian, Y., et al. (2021). Weakly-supervised video anomaly detection with robust temporal feature magnitude learning. ICCV, 4975-4986.

Wang, X., et al. (2018). Repulsion loss: Detecting pedestrians in a crowd. CVPR, 7774-7783.

Wilhelmsen, C. A., & Dixon, R. A. (2016). Identifying indicators for engineering design outcome. Journal of Technology Education, 27(2), 57-77.

Zhu, X., et al. (2017). Flow-guided feature aggregation for video object detection. ICCV, 408-417.

Zou, Z., et al. (2023). Object detection in 20 years: A survey. Proceedings of the IEEE, 111(3), 257-348.

## Appendix: Mathematical Formulations

Each component is normalized to [0, 1] where higher values indicate greater complexity.

**Crowd Density Index (CDI)**

$$CDI = 0.4 * D_{norm} + 0.6 * O_{avg}$$

**Where:** D_norm = detection count / 50; O_avg = mean pairwise IoU

**Camera Quality Index (CQI)**

$$CQI = 0.4 * B + 0.3 * N + 0.3 * R$$

**Where:** B = blur (Laplacian variance); N = noise (FFT ratio); R = resolution penalty

**Motion Randomness Index (MRI)**

$$MRI = 0.5 * M_{norm} + 0.5 * E_{dir}$$

**Where:** M_norm = normalized flow magnitude; E_dir = directional entropy

**Luminance Variability Index (LVI)**

$$LVI = 0.3 * C + 0.3 * V + 0.4 * L$$

**Where:** C = luminance std/128; V = contrast ratio; L = low-light penalty

**Temporal Outlier Deviation Index (TODI)**

$$TODI = |F - mean| / (3 * std)$$

**Where:** Clamped to [0,1]; TODI = 1.0 for static images

**Campus Complexity Index (CCI)**

$$CCI = 0.2 * CDI + 0.2 * CQI + 0.2 * MRI + 0.2 * LVI + 0.2 * TODI$$

**Where:** Equal weights; CCI in [0,1]