Safety Helmet Detection Based on Improved YOLOv8n-SLIM-CA

Ravina Suresh Edake[1], Dr. Niranjan Tukaram Kulkarni[2]

[1]Student, Alamuri Ratnamala Institute of Engineering and Technology (ARMIET), Maharashtra, India
[2]Associate Professor, New Horizon Institute of Technology and Management (NHITM)

Abstract - *Ensuring worker safety in construction and industrial settings always comes down to proper monitoring, and making sure everyone actually wears their safety helmets is a huge part of that. But traditional manual supervision is not only exhausting for the people stuck watching footage all day; it's also error-prone and just not practical when it comes to large-scale, round-the-clock monitoring. Deep learning and object detection have changed the game for real-time safety monitoring. YOLO's speed and accuracy set a standard, but tracking small, far-off, or partly hidden helmets on chaotic work sites? That still remains a significant challenge for existing systems. This research introduces an upgraded safety helmet detection system built on a customized YOLOv8 base called YOLOv8n-SLIM-CA. Here's what makes it tick: we use mosaic data augmentation to boost the network's sensitivity to small objects and make it generalize better, even if the work site always looks different. Coordinate Attention (CA) gets baked into the backbone network, sharpening the model's focus on helmet regions and dampening visual distractions in the background. We slim down the neck architecture for lean, speedy multi-scale feature fusion ideal for lower-power, real-time applications. On top of that, we drop in a dedicated small target detection layer to crank up accuracy for far-off workers; you spot more helmets, even at a distance or in a crowd. We test the model with all the standard metrics: precision, recall, and mean Average Precision (mAP). Results show pretty clearly detection improves for helmets, even in complicated, high-clutter environments, all while keeping things fast enough for live deployment. Since this build has a tight computational footprint, it plays well with edge devices and embedded hardware. With YOLOv8n-SLIM-CA, we offer a smarter, lighter way to automate safety checks, so compliance goes up and risk goes down.*

Key Words: Safety Helmet Detection, YOLOv8, Deep Learning, Object Detection, Coordinate Attention, Small Target Detection, Computer Vision, Real-Time Monitoring, Industrial Safety, Edge Computing

1. INTRODUCTION

Worker safety for industries like construction and mining usually centers around rigid rules for wearing protective gear especially helmets. Head injuries are no joke; they're a leading cause of serious accidents mostly because people cut corners or the boss misses someone not following protocol. The standard way? Watching security videos manually. That just sets up a bunch of predictable problems: supervisors get tired, responses to violations are slow, and you're stuck with subjective error-prone assessments.

As artificial intelligence pushes deeper into safety solutions, computer vision models especially deep learning-based object detectors are making it possible to keep eyes on every worker at all times, without human fatigue. The YOLO architecture has built a reputation for balancing high accuracy and speed, and the latest iteration, YOLOv8, adds even better feature extraction and fast inference. Even then, recognizing helmets out in the wild is tricky.

Open construction zones look nothing like clean test images there are messy backgrounds, weird lighting, and too many people in frame. Helmets end up small or blurred by distance, and overlapping workers or equipment hide them further. All this means YOLO models (and their competitors) often call out helmets where there are none, or worse, completely miss real ones.

To solve these issues, simply stacking more layers or making the network deeper doesn't cut it models grow too bulky and slow for edge devices. Instead, injecting smarter modules like channel and spatial attention, plus streamlining multi-scale feature fusion, can make a model sharper and more robust without bogging down processing.

That's where YOLOv8n-SLIM-CA comes in. By adding mosaic augmentation, coordinate attention, a slimmed-down neck, and a small-object detection layer, this version nails down the details that legacy systems miss and runs light enough for real-time, on-site deployment.

2. LITERATURE REVIEW

N. Fatima and her team (2025) [1] bolstered YOLOv8n with Coordinate Attention and a Slim-Neck structure. By highlighting meaningful image regions and discarding noisy features, they made the detection of tiny, far-away helmets more reliable. Slim-Neck stitched together multi-scale features without fluff, but, admittedly, the upgrades added to model complexity possibly making lightweight deployment more difficult.

Then there's the FGP-YOLOv8n model from L. Zhang (2025) [2], which swapped out the usual backbone for FasterNet and slotted in lightweight attention layers. The result? The model not only trimmed down resource demands but still lifted mAP. The catch is, by altering the backbone so drastically, you sometimes sacrifice flexibility for other object detection tasks.

C. Song and Y. Li (2025) [3] tuned YOLOv8 with beefed-up feature aggregation and better upsampling, letting data flow smoothly across scales. These improvements let the model dig deeper for spatial and semantic cues, even when objects overlapped. Still, feature aggregation upped the average inference times, which isn't ideal for rapid-response settings.

S. J. Li's MH-YOLO (2024) [4] built on YOLOv8 by weaving in CBAM, targeting improved feature extraction. Further, with special modules for tiny, far-off, or overlapping helmets detection rates soared. But with all these add-ons, runtime and edge performance did take a hit.

X. Wu and colleagues (2023) [5] offered CC-YOLOv8, using the C2fcc block for extraction and EMA for stronger object focus. While detection accuracy went up, their evaluation didn't put enough emphasis on model speed and deployment on low-power gear leaving some open questions about real-world viability.

3. PROBLEM STATEMENT

Despite leaps in deep learning for vision, pinpointing helmets reliably on busy construction sites is far from solved. Models like YOLO stumble when helmets are small, far off, or hidden even for the latest releases. Most helmets cover few pixels in surveillance footage, making it tricky for networks to extract clean features, which translates to more false negatives.

What makes matters worse? Complicated backgrounds crammed with machinery, scaffolding, and changing patterns confuse the system, prompting mistakes while identifying or classifying helmets. Factor in a wide range of weird lighting shadows one moment, harsh reflections the next and even the top models lose steam.

And then there's the accuracy vs. speed trade-off. Most approaches dial up accuracy only by making networks deeper or tossing in new layers, which is a problem when you want to run these models on edge devices and not in a server room with unlimited compute.

Finally, without specific modules for ignoring background clutter and spotlighting helmet regions, false positives spike especially in crowded frames.

So, you get that the real need is obvious: a model that can find safety helmets fast, accurately, and efficiently. It should work with small, hard-to-see helmets, handle noisy backgrounds, and be lean enough to run on-site in real time.

4. METHODOLOGY

We're building our solution around a souped-up YOLOv8 architecture tailored to hunt for helmets, even in the messiest environments, and without wasting resources. Everything starts at the source: we process video or images from site cameras, prepping the data through resizing, normalization, and aggressive augmentation.

Fig -1: Methodology for Safety Helmet Detection

Mosaic augmentation sits at the core of our pipeline. By mixing and matching images to form new training collages, we force the network to see more small helmets in more scenarios. This simple trick pays dividends in real-world variation.

For feature extraction, we retrofit the backbone with Coordinate Attention. Now the network can figure out both the "where" and the "what," focusing energy on helmets and dialling down distractions without losing the big picture.

To keep things lightweight, we use a Slim-Neck. It ties together information from different scales, keeping the best of each layer without bloating computation. Efficiency stays high, while the model doesn't miss subtle, small details.

We install a dedicated detection layer for tiny targets. This means the system is way more likely to pick up workers at a distance or in dense crowds.

Once the features are processed, the detection head does its thing: it draws bounding boxes, hands out "helmet" or "no helmet" tags, and ranks its confidence. With Non-Maximum Suppression (NMS), we clean up duplicates, leaving only the most probable detections.

To sum up results, we use precision, recall, and mAP standard tools, making it clear where the model wins or needs work.

5. WORKING

Here's how it plays out on site: cameras feed video to the system. Each frame is tuned for size and quality, then piped through the network.

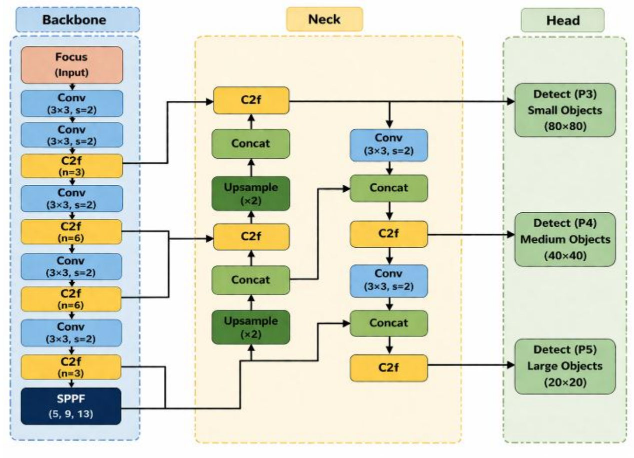

Fig -2: Working for Safety Helmet Detection

YOLOv8's backbone, charged with Coordinate Attention, breaks the image into features and selectively shines a spotlight on helmet-like areas. These features pass through the efficient Slim-Neck, getting combined and refined across scales. The special small-target layer catches those easy-to-miss helmets at the back of the shot or in a crowd.

After this, the detection head marks up bounding boxes around helmets and gives each a confidence rating. NMS cleans up overlap, so you're left with sharp, non-redundant predictions.

All this happens live. If the system notices someone without a helmet, instant alerts can be triggered, prompting site-wide or targeted warnings. Since the model is lightweight, you can run it on a standard edge device keeping delays minimal and resource costs down. Building our solution around a souped-up YOLOv8 architecture tailored

6. CONCLUSION

All combined YOLOv8n-SLIM-CA steps up to the plate for helmet detection on busy, unpredictable sites. With clever data handling, smarter feature focusing, streamlined fusion, and targeted small-object modules, you're looking at a system that does what prior solutions struggled with: catching helmets under challenging conditions, without lugging around heavyweight hardware.

This approach balances speed and smarts you can actually deploy this on-site, not just in an ideal lab. Cameras keep rolling, safety managers get reliable alerts, and you don't need to babysit the system or do manual spot checks. The result is simple: lower accident risk and stronger workplace compliance, all through automation.

7. FUTURE SCOPE

Next steps look promising. Tackling the hardest cases really low light, thick occlusions could push the model even further, perhaps by shifting toward transformer-based architectures that excel at context understanding.

Integrating the solution into larger IoT safety frameworks opens up centralized control and analytics across multiple sites. Beyond helmets, it makes sense to train the model for vests, gloves, and other safety wear, creating fuller protection suites.

Hardware-wise, refining the model via quantization and leaning into hardware acceleration (using chips tuned for AI inferencing) will make rollouts even smoother and cheaper. Adaptive or continual learning, where the model keeps training after deployment, could drive cross-site generalization and performance up as conditions change.

REFERENCES

[1] N. Fatima et al., "Safety Helmet Detection Based on Improved YOLOv8," 2025.

[2] L. Zhang et al., "An Improved Lightweight Safety Helmet Detection Algorithm for YOLOv8," 2025.

[3] C. Song and Y. Li, "YOLOv8 Algorithm for Safety Helmet Detection in Complex Environments," 2025.

[4] S. J. Li et al., "Real-time Helmet Detection Based on Improved YOLOv8," 2024.

[5] X. Wu et al., "Improved Safety Helmet Detection Using YOLOv8," 2023.