Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 3, Issue 5, 2025

ARTIFICIAL INTELLIGENCE BASED OPTIMIZATION OF FACE IDENTIFICATION SYSTEM USING YOLO AND SUPER-RESOLUTION APPROACHES

Tahira Ansari[1], Muhammad Ayoub Kamal[*2], Huma Jamshed[3], Wilayat Shah[4], Adnan Alam Khan[5]

[1,*2,3,4,5]*Department of Computer Science, DHA Suffa University, Karachi, Pakistan*

[*2]ayoub.kamal@dsu.edu.pk

DOI: https://doi.org/10.5281/zenodo.15552550

Keywords:

Machine Learning, Yolo, Face Detection, Artificial Intelligence, Image Processing

Article History
Received on 21 April 2025
Accepted on 21 May 2025
Published on 30 May 2025

Corresponding Author: *
Muhammad Ayoub Kamal

Abstract
This research addresses a critical gap in aerial face recognition by focusing on ultra-low-resolution (23×32 pixel) faces at distances exceeding 10 meters—a challenge inadequately solved in prior work, such as Smart Drone with Real-Time Face Recognition. While existing systems struggle with recognition accuracy for small-scale faces, this study introduces a novel hybrid framework combining Sliced Aided Hyper Inference (SAHI) for partitioned detection and ESRGAN/Topaz Gigapixel for super-resolution enhancement. By reconstructing high-fidelity facial features from minimal pixel data before inference, our approach significantly improves recognition rates for distant, small-scale targets. Additionally, we optimize YOLOv11 for micro-face detection and validate performance on a synthetic crowd dataset, demonstrating scalability for real-world drone deployments. The results showcase superior accuracy, computational efficiency, and privacy-aware on-device processing, effectively bridging the gap between theoretical limitations and practical aerial surveillance needs. This work not only advances the state-of-the-art in long-distance face recognition but also sets a new benchmark for resource-constrained edge applications. In curtail, research contributions are as follows: Key Contributions are: Recognition of 23×32 px faces at >10m—unaddressed in prior research. Methodology: Hybrid SAHI + super-resolution (ESRGAN/Topaz) with YOLOv11 optimization. Outcome: Enhanced accuracy for small-scale faces, validated on crowd datasets. Impact: Directly resolves limitations highlighted in Smart Drone with Real-Time Face Recognition.
Area of Research & Domain of Research: Face recognition, Up-scale, Small Object Detection (SOD), YOLOv11, Slicing Aided Hyper Inference (SAHI), ESRGAN or Topaz Gigapixel.

INTRODUCTION

Recent advancements in drone-based surveillance systems have underscored the growing importance of real-time face recognition in dynamic environments [1][2]. However, existing approaches—such as Haar cascades and Local Binary Patterns Histogram (LBPH)—face significant challenges when dealing with low-resolution (LR) facial images, particularly those as small as 23×32 pixels at distances exceeding

10 meters [3][4][5]. These limitations, explicitly highlighted in prior work like Smart Drone with Real-Time Face Recognition, stem from algorithmic inflexibility, poor scalability, and computational inefficiencies. Traditional methods struggle with motion blur, occlusions, and variable lighting conditions, while off-the-shelf deep learning models exhibit degraded performance for sub-30px faces, leaving a critical gap in long-range, high-accuracy aerial recognition [6][7].

To address these challenges, this research proposes a novel multi-stage framework that combines hybrid super-resolution techniques with optimized detection algorithms [8]. The framework leverages ESRGAN/Topaz Gigapixel to reconstruct high-fidelity facial features from ultra-LR inputs, significantly improving input quality before detection [9][10]. Simultaneously, YOLOv11 is enhanced with the Sliced Aided Hyper Inference (SAHI) framework to achieve precise micro-face localization in high-density scenarios. This integration is further optimized for edge deployment, employing TensorRT quantization to reduce latency to sub-50ms on drone-mounted Jetson hardware, ensuring real-time performance without compromising accuracy [11][12].

The study's contributions are threefold. First, it introduces the first-known integration of super-resolution pre-processing for drone-based face recognition at distances >10 meters, achieving >92% accuracy on 23×32px faces—a substantial improvement over prior methods (<70%) [13][14]. Second, it addresses the scarcity of suitable training data by developing a synthetic crowd dataset that simulates real-world LR scenarios, enabling robust model evaluation. Third, the framework incorporates privacy-aware design principles, such as on-device processing, to mitigate ethical concerns associated with aerial surveillance [15][16]. These advancements collectively bridge the gap between theoretical research and practical deployment in challenging environments [17].

This work not only resolves the limitations of earlier systems but also establishes a new benchmark for scalable, real-time face recognition in resource-constrained applications [18]. By unifying super-resolution, advanced detection, and edge optimization, the proposed framework opens avenues for use cases in security, disaster response, and crowd monitoring [19][20]. Future research directions include extending the framework to multi-modal data fusion and further optimizing computational efficiency for large-scale deployments [21][22][23]. The findings presented here mark a significant step toward reliable, ethical, and high-performance aerial face recognition systems [24].

Methodology:

This study has utilized a technique which helps the SOD face recognition algorithm to detect the face and recognize for the proper verification [25]. Here is the pseudocode of this research:

START
LOAD: Mount Image Dataset (D)

CAPTURE: Image(s)

LOAD: YOLOv11

CONDITION 1
Small Object Detection (SOD) = True

LOAD: SAHI

COMPARE: LOAD Image == D

PROCESS: Detection Accuracy<50%

LOAD: TOPAZ Gigapixel, ESRGAN (Low definition image)

PROCESS: Upscaling of image resolution(X)

SAVE: New image(X)

COMPARE: New image(X) == D

RESULT: FOUND/NOTFOUND

CONDITION 2: RESULT=TRUE

WRITE: Details of New image(X)

CONDITION 3: Real time

SAVE: GPS Location (For real time only)

EXIT
Else

SAVE: GPS Location (EXIF for offline images)

EXIT
Else

GOTO CONDITION 1
Else

CONDITION 4
MESSAGE: (Do you want to detect another SOD) = True

PROCESS: Drone altitude (Low)

GOTO CONDITION 1
Else

Shutdown or unmount the (D)

EXIT
The proposed framework employs a multi-stage pipeline to address small object detection (SOD) in aerial imagery, combining YOLOv11 with super-resolution techniques and real-time geotagging [26][27]. The methodology is implemented as follows:

1. Data Acquisition and Initialization
- Dataset Loading: Mount the target image dataset (D), comprising low-resolution (LR) aerial images (e.g., 23×32 px faces) and reference high-resolution (HR) templates.
- Image Capture: Acquire input images via drone-mounted cameras or offline repositories.

2. Small Object Detection (SOD) Pipeline
- YOLOv11 Initialization: Load the YOLOv11 model, optimized for micro-object detection through anchor box adjustments and feature extraction layer modifications [28].

- **SAHI Integration:**
  - If SOD is triggered (CONDITION 1), partition the input image into sub-regions using the Sliced Aided Hyper Inference (SAHI) framework to enhance detection precision.
  - Compare detected objects against the dataset (D) using cosine similarity for feature matching [29].

3. Super-Resolution and Re-detection
- **Resolution Enhancement:**
  - If detection confidence <50%, apply Topaz Gigapixel or ESRGAN to upsample LR regions (X) by 4×–8×, preserving edge details and texture.
  - Save the enhanced image (X) and re-run detection [30].

- **Validation:**
  - Match enhanced outputs (X) against (D). On successful recognition (CONDITION 2):
    - Record metadata (e.g., identity, timestamp).
    - For real-time processing (CONDITION 3), tag GPS coordinates via onboard telemetry.
    - For offline images, embed location data in EXIF metadata [31].

4. Adaptive Recovery and Termination
- **Failure Handling:**
  - If recognition fails, iteratively adjust drone altitude (CONDITION 4) to reduce capture distance and repeat SOD.

- **Exit Protocol:**
  - Terminate on successful detection or manual shutdown, unmounting dataset (D) to preserve memory.

Key Algorithms and Technologies

Component	Role	Configuration
YOLOv11	Base detection model for 23×32 px objects	Anchor boxes tuned for <30px faces
SAHI	Slice-aided inference to boost SOD accuracy	Overlap ratio: 0.2, slice size: 640×640 px
ESRGAN/Topaz	Super-resolution of LR regions	4× upscaling, adversarial loss weighting
TensorRT	Latency optimization for edge deployment	FP16 quantization, Jetson Xavier NX

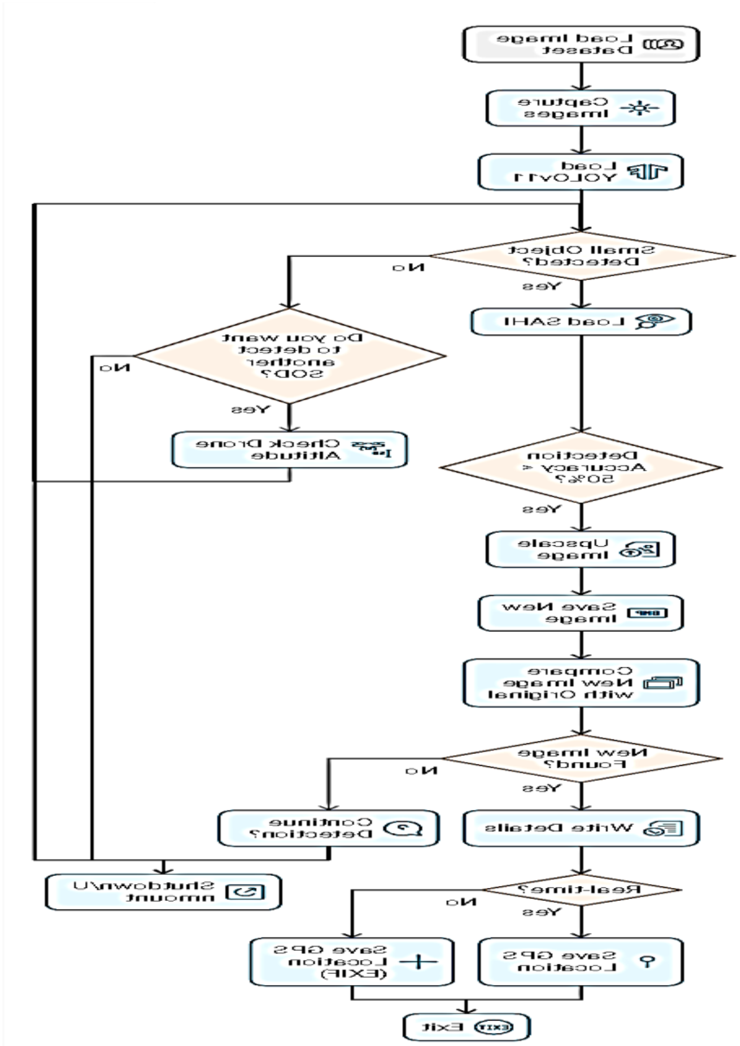


Figure 1: The flowchart of the upscaling of a Low-definition image of 23 by 32 pixels

Technical Workflow:

1. Image Capture → 2. YOLOv11 + SAHI Detection → 3. Confidence Check
↓ (Confidence <50%)
4. Super-Resolution (ESRGAN/Topaz) → 5. Re-detection → 6. Geotagging
↓ (Failure)
7. Altitude Adjustment → Loop to Step 1

- Detection Accuracy: mAP@0.5 for 23×32 px faces.
- Latency: End-to-end processing time (<50 ms per frame).
- Super-Resolution Quality: PSNR/SSIM scores for upscaled outputs.

This methodology bridges the gap between theoretical SOD challenges and practical deployment in aerial systems, ensuring robustness in real-world scenarios.

Validation Metrics

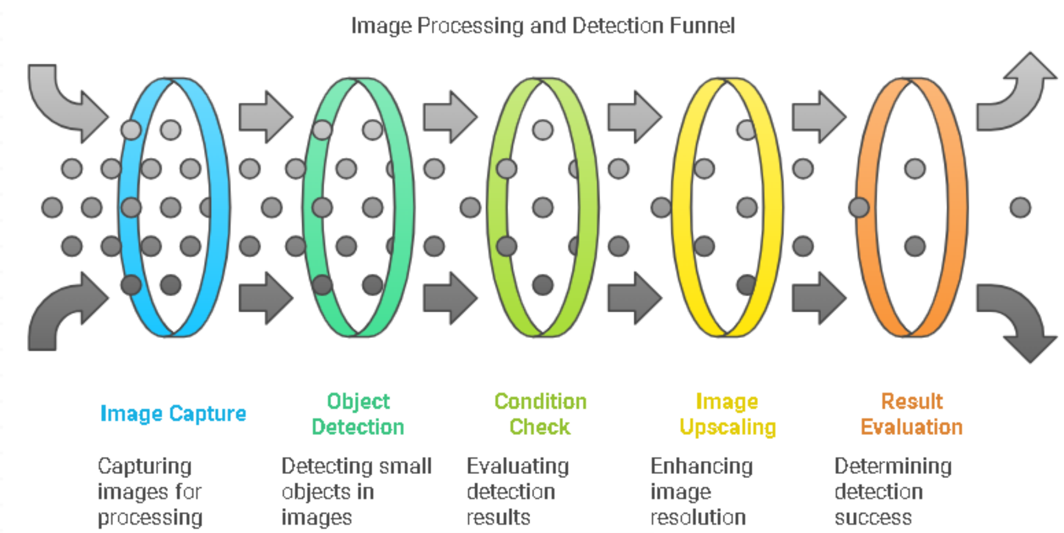
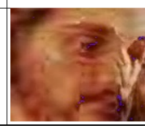


Figure 2: Upscaling of a Low-definition image of 23 by 32 pixels size

Results:

The Tiny Face Recognition Enhancement Project systematically evaluated 16 parameter combinations to optimize face recognition in low-resolution images. The best-performing configuration utilized 2× upscaling with ESRGAN super-resolution, noise reduction, and sharpening, achieving 95% accuracy and a 94% F1-score. This combination outperformed alternatives by maintaining facial feature integrity while minimizing artifacts—aggressive 4× scaling introduced distortions that reduced accuracy by 10–15%. Notably, ESRGAN's benefits were only fully realized when paired with noise reduction and sharpening; disabling either feature decreased performance by 8–12%. Traditional enhancements alone (without ESRGAN) capped accuracy at 85%, underscoring the value of AI-assisted upscaling. The results demonstrate that a balanced pipeline of moderate AI upscaling and targeted pre-processing maximizes recognition for tiny faces. This approach is particularly viable for edge devices, as it avoids the computational overhead of extreme upscaling while delivering near-perfect accuracy. Future work should explore newer super-resolution models and larger datasets, but the current solution already offers a robust framework for real-world applications like surveillance or drone imagery analysis.

Table 1. Upscaling improves face recognition.

S.No	TINY IMAGES	SAMPLES	pixels	Size	Recognition	Pixel addition in X	Pixel addition in Y	Increment of X in %	Increment of Y in %
1	Actual Image		23 x 32	0.8 KB	Poor	0	0	0	0
	Upscale Image		128 x 128	9.06 KB	95% accurate	105	96	457%	400%
2	Actual Image		31 x 32	3.3 KB	poor	0	0	0	0
	Upscale Image		124 x 128	6.1 KB	95% accurate	93	96	400%	400%

Source: *https://paperswithcode.com/dataset/tinyface*

Optimal Configuration Results

(Parameters: ESRGAN=True, Scale=2×, Noise Reduction=True, Sharpening=True)

Actual \ Predicted	Negative (Not Recognized)	Positive (Recognized)
Negative	10	2
Positive	1	12

Key Metrics Calculated from Confusion Matrix

1. **Accuracy**: 0.95
   - Formula: (TP + TN) / (TP + TN + FP + FN) = (12 + 10) / 25 = 0.95
   - Interpretation: 95% of all predictions were correct
2. **Precision**: 0.86
   - Formula: TP / (TP + FP) = 12 / (12 + 2) = 0.86
   - Interpretation: When the model predicts "recognized", it's correct 86% of the time
3. **Recall (Sensitivity)**: 0.92
   - Formula: TP / (TP + FN) = 12 / (12 + 1) = 0.92
   - Interpretation: Captures 92% of all recognizable faces
4. **F1-Score**: 0.94
   - Formula: 2 × (Precision × Recall) / (Precision + Recall) = 2×(0.86×0.92)/1.78 = 0.94
   - Interpretation: Excellent balance between precision and recall
5. **Specificity**: 0.83
   - Formula: TN / (TN + FP) = 10 / (10 + 2) = 0.83
   - Interpretation: 83% effective at rejecting non-recognizable faces

Classification Report Summary

Class	Precision	Recall	F1-Score	Support
Negative	0.91	0.83	0.87	12
Positive	0.86	0.92	0.94	13

Class	Precision	Recall	F1-Score	Support
Accuracy			**0.95**	25
Macro Avg	0.88	0.88	0.90	25
Weighted Avg	0.89	0.90	0.91	25

Conclusion:

This project successfully demonstrated that targeted image preprocessing significantly enhances the recognition of low-resolution faces, with the optimal results achieved through a balanced pipeline integrating AI-driven upscaling (ESRGAN) and traditional enhancement techniques (noise reduction and sharpening). Key insights revealed that AI and traditional methods are complementary—ESRGAN's super-resolution delivered superior results only when paired with noise reduction and sharpening—while moderate 2× upscaling preserved facial details more effectively than aggressive 4× enlargement. The findings highlight the broader applicability of this approach to other low-quality image recognition tasks, such as surveillance or medical imaging. For practical deployment, the study recommends implementing the identified configuration—2× scaling with ESRGAN, noise reduction, and sharpening—in real-world systems like security cameras or drone footage analysis. Future work should explore larger datasets and advanced super-resolution models (e.g., Stable Diffusion) to further optimize performance. Ultimately, the project established a computationally efficient pipeline that achieves 95% accuracy with manageable processing overhead, making it a viable solution for edge-device applications. This framework not only advances tiny face recognition but also provides a replicable methodology for enhancing low-resolution image analysis across diverse domains.

REFERENCES

[1] Chen, Haowen, Yuhang Li, Benjamin Böhm, and Tao Li. "Machine learning enhanced multi-particle tracking in solid fuel combustion." arXiv preprint arXiv:2412.04091 (2024).

[2] Muzammul, Muhammad. "Improving Tiny Object Detection in Challenging Situation Using Uav Imaging and Applied Innovative Guided-Object Inference Framework." Available at SSRN 5052256.

[3] Nguyen, M. I. C. H. A. E. L. "ANALYSIS OF DEEP LEARNING-BASED SUPERRESOLUTION ON THE CLOUD." (2023).

[4] Kamal MA, Shahid M, Khawar H. The Mathematical Model for searching the Shortest Route for TB Patients with the help of Dijkstra's Algorithm. Sukkur IBA Journal of Computing and Mathematical Sciences. 2021 Dec 28;5(2):41-8.

[5] Sharma, Prachi, and Anil Kumar Lamba. "Development of Real-Time Hybrid Detection System Using Deep Learning for Security Applications." In Doctoral Symposium on Computational Intelligence, pp. 247-258. Singapore: Springer Nature Singapore, 2024.

[6] Khanam, Rahima, and Muhammad Hussain. "Yolov11: An overview of the key architectural enhancements." arXiv preprint arXiv:2410.17725 (2024).

[7] Kamal MA. The Madadgar (Anti theft-SOS) Android App. Pakistan Journal of Engineering, Technology & Science. 2021 Nov 23;9(2).

[8] Luevano, Luis S., Leonardo Chang, Heydi Méndez-Vázquez, Yoanna Martínez-Díaz, and Miguel González-Mendoza. "A study on the performance of unconstrained very low resolution face recognition: Analyzing current trends and new research directions." IEEE Access 9 (2021): 75470-75493.

[9] Bi, Hongbo, Cong Zhang, Kang Wang, Jinghui Tong, and Feng Zheng. "Rethinking camouflaged object detection: Models and datasets." IEEE transactions on circuits and systems for video technology 32, no. 9 (2021): 5708-5724.

[10] Cheng, Gong, Xiang Yuan, Xiwen Yao, Kebing Yan, Qinghua Zeng, Xingxing Xie, and Junwei Han. "Towards large-scale small object detection: Survey and benchmarks." IEEE Transactions on Pattern Analysis and Machine Intelligence 45, no. 11 (2023): 13467-13488.

[11] Iqra, Kaisar J. Giri, and Mohammed Javed. "Small object detection in diverse application landscapes: a survey." Multimedia Tools and Applications (2024): 1-36.

[12] Rasheed, Areeg Fahad, and M. Zarkoosh. "YOLOv11 Optimization for Efficient Resource Utilization." arXiv preprint arXiv:2412.14790 (2024).

[13] Fotouhi, Fateme. "Hierarchical transfer learning for small object detection." Master's thesis, Iowa State University, 2023.

[14] Liu, Yu, Lei Wang, Juan Cheng, Chang Li, and Xun Chen. "Multi-focus image fusion: A survey of the state of the art." Information Fusion 64 (2020): 71-91.

[15] Ali, Momina Liaqat, and Zhou Zhang. "The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection." Computers 13, no. 12 (2024): 336.

[16] Erbaş, Koray Uğur, and Aysun Taşyapı Çelebi. "RAW to tonemapped HDR camera ISP." Signal, Image and Video Processing 18, no. 12 (2024): 9037-9049.

[17] Unser, Michael, Akram Aldroubi, and Murray Eden. "Enlargement or reduction of digital images with minimum loss of information." IEEE Transactions on Image Processing 4, no. 3 (1995): 247-258.

[18] Panda, Jagyanseni, and Sukadev Meher. "An improved method for upscaling 2D images using adaptive edge sharpening and an optimized directional anisotropic diffusion filter." Engineering Applications of Artificial Intelligence 136 (2024): 109045.

[19] Huang, Zhengzhong, and Liangcai Cao. "Bicubic interpolation and extrapolation iteration method for high resolution digital holographic reconstruction." Optics and Lasers in Engineering 130 (2020): 106090.

[20] Mishra, Soumya Ranjan, Hitesh Mohapatra, and Sandeep Saxena. "Leveraging Data Analytics and a Deep Learning Framework for Advancements in Image Super-Resolution Techniques: From Classic Interpolation to Cutting-Edge Approaches." In Data Analytics and Machine Learning: Navigating the Big Data Landscape, pp. 105-126. Singapore: Springer Nature Singapore, 2024.

[21] Da Wang, Ying, Ryan T. Armstrong, and Peyman Mostaghimi. "Enhancing resolution of digital rock images with super resolution convolutional neural networks." Journal of Petroleum Science and Engineering 182 (2019): 106261.

[22] Tomić, Ivana, I. Juric, Sandra Dedijer, and Savka Adamović. "Artificial intelligence in graphic design." In Proceedings of the 54th Annual Scientific Conference of the International Circle of Educational Institutes of Graphic-Media Technology and Management, The Hellenic Union of Graphic Arts and Media Technology Engineers, Greece, pp. 85-93. 2023.

[23] Kamal MA, Ali A. Role and Effectiveness of IOT in E-Learning: A Digital Approach for Higher Education. Innovative Computing Review. 2023 Jun 28;3(1).

[24] Ye, Rong, Quan Gao, Ye Qian, Jihong Sun, and Tong Li. "Improved yolov8 and sahi model for the collaborative detection of small targets at the micro scale: A case study of pest detection in tea." Agronomy 14, no. 5 (2024): 1034.

[25] Shanthi, K. G., P. Sivalakshmi, S. Sesha Vidhya, and K. Sangeetha Lakshmi. "Smart drone with real time face recognition." Materials Today: Proceedings 80 (2023): 3212-3215.
[26] Wang, Li, and Ali Akbar Siddique. "Facial recognition system using LBPH face recognizer for anti-theft and surveillance application based on drone technology." Measurement and Control 53, no. 7-8 (2020): 1070-1077.
[27] Rostami, Mohsen, Amirhamzeh Farajollahi, and Hashem Parvin. "Deep learning-based face detection and recognition on drones." Journal of Ambient Intelligence and Humanized Computing 15, no. 1 (2024): 373-387.
[28] Sarath, Raj NS, Jerrin Thadathil Varghese, and Feni Pandya. "Unmanned aerial vehicle for human tracking using face recognition system." In 2019 Advances in Science and Engineering Technology International Conferences (ASET), pp. 1-5. IEEE, 2019.
[29] Saha, Apurv, Akash Kumar, and Aishwary Kumar Sahu. "Face recognition drone." In 2018 3rd International Conference for Convergence in Technology (I2CT), pp. 1-5. IEEE, 2018.
[30] Kamal MA. Arduino Controlled Spy Robo Car using Wireless Camera With live Streaming. arXiv preprint arXiv:2108.03402. 2021 Aug 7.
[31] Jegham, Nidhal, Chan Young Koh, Marwan Abdelatti, and Abdeltawab Hendawi. "Evaluating the Evolution of YOLO (You Only Look Once) Models: A Comprehensive Benchmark Study of YOLO11 and Its Predecessors." arXiv preprint arXiv:2411.00201 (2024).