



COLLEGE OF CREATIVE INNOVATION

Programme
Master of Software Engineering
180 Credits

Course

MSE806: Intelligent Transportation Systems
(15 credits)

Assessment 2

Group Project Assignment

Student's Name	Student's ID
Alfredo Jr. Albis Torreña	270581914
Eunseok Choi	270597067
Fabricio Mello Mulato	270516212
Yan Huang	270466925

Supervisor's Name: Mukesh Mishra

Date: July 2025

Table of Contents

1	Introduction	3
2	Project Proposal	4
2.1	Project Title	4
2.2	Objectives	4
2.3	Technological Innovation	5
2.4	Ethical Considerations	6
3	Literature Review	6
3.1	Fundamental Concepts	6
3.2	Related Work	8
3.3	Limitations and Gaps in the Literature	9
4	Methodology	11
4.1	Approach	11
4.2	System Design and Architecture	12
4.3	Implementation Steps	14
5	Results and Analysis	17
5.1	Software Features	17
5.2	Performance Metrics	17
5.3	Security Assessment	20
6	Discussion	21
6.1	Contributions and Challenges	21
6.2	Future Enhancements	22
7	Conclusion and Reflection	23

References	25
Appendix: Training Results and Visualisations	29
Appendix B: Complete Project Source Code	31

1 Introduction

Despite significant improvements in vehicle safety systems, many accidents still happen in driveways, which are often considered low-risk areas. In New Zealand, an average of five children are killed or injured each year in such incidents, with toddlers accounting for nearly 90% of the cases (Otago Daily Times Online News, [2025](#)). A similar pattern exists in the United States, where two children die each week due to vehicle backovers, highlighting a widespread issue (KidsAndCars.org, [2023](#)).

These accidents typically occur when a driver reverses without noticing a child nearby. Children under five are most vulnerable, and more than 70% of the time, the driver is a close family member. The combination of limited visibility and unpredictable child behaviour increases the likelihood of severe outcomes in these situations.

To address this gap in current safety technology, the project proposes an intelligent driveway detection system. It uses advanced computer vision and machine learning to identify small or partially hidden objects in different conditions. Integrated with existing camera systems, it provides visual and audible alerts in real time, with the potential for automatic braking. The system is designed to offer timely warnings while reducing false alarms, aiming to prevent injuries and fatalities where existing solutions fail.

2 Project Proposal

2.1 Project Title

"AI Driven Driveway Safety System: Preventing Child Accidents with Computer Vision"

2.2 Objectives

The primary objectives of this project are as follows:

1. To design and prototype a safety system using computer vision and machine learning to detect people, children, and pets near vehicles, especially in driveways.
2. To develop a user-friendly vehicle dashboard interface delivering visual, auditory, or haptic feedback, with context-aware detection that adjusts sensitivity and alert thresholds based on the vehicle's status, reducing false positives while ensuring critical risks are flagged.
3. To prevent accidents by providing timely and clear alerts that reduce injuries and fatalities caused by limited driver visibility during vehicle manoeuvres in driveways.

This project addresses a specific scenario often overlooked by traditional Advanced Driver Assistance Systems (ADAS), including Parking Assistance and Reverse Warning features, which may fail to detect small or partially hidden children in residential driveways.

These objectives contribute to improving vehicle safety by focusing on a specific but important scenario often neglected by traditional Advanced Driver Assistance Systems (ADAS). The concept of ADAS and its role within the context of Intelligent Transportation Systems will be addressed in Section 3.

2.3 Technological Innovation

The proposed system combines advanced technologies to provide a reliable and responsive safety solution for driveways. The key components are as follows:

1. **Contextual Behaviour Detection:** Detection settings will adapt to the vehicle's movement, increasing sensitivity at low speeds during parking to reduce missed detections and unnecessary alerts.
2. **Sensor Fusion (Optional):** Adding sensors such as infrared and ultrasonic will improve detection in poor light or blocked views, making the system more reliable.
3. **Edge Computing:** Processing will happen locally on devices like NVIDIA Jetson Nano to ensure quick responses, protect privacy, and keep costs down.
4. **Dashboard Integration:** Alerts will use visuals, sounds, and vibrations to notify drivers clearly, with possible automatic braking if warnings are ignored.

This technological approach aims to address the unique safety challenges of residential environments, offering a targeted and proactive solution beyond conventional ADAS.

2.4 Ethical Considerations

The system tackles key ethical and safety issues in computer vision involving human data. To comply with GDPR, all video is processed locally without cloud storage, protecting user privacy. Datasets like COCO (Common Objects in Context) often lack informed consent, raising concerns about bias, especially in detecting children (Brandao, 2019; Tahir, 2024). This was addressed by retraining the model with diverse images and adding a dedicated class for kids. Another issue to consider is blurring the faces of people and children in training, validation, and testing images. It is important to note that users can adjust sensitivity settings, and the interface reinforces that responsibility remains with the driver.

3 Literature Review

3.1 Fundamental Concepts

Intelligent Transportation Systems (ITS) cover a wide range of solutions that combine information and communication technologies within the transport sector. The main goal is to improve economic, social, and energy outcomes by making the interaction between vehicles, infrastructure, and users more efficient across all types of transport (Perallos et al., 2016).

This project aligns with the field of ITS, where computer vision is used for various tasks such as positioning and controlling autonomous vehicles and monitoring traffic

flow (Yuan et al., 2019). According Akshayaa and Nithin (2021) and Yuan et al. (2019), pedestrian detection is a key concern in this field, aimed at improving road user safety and preventing accidents.

Key areas within Intelligent Transportation Systems (ITS) include Advanced Driver Assistance Systems (ADAS) (Miani et al., 2022), machine learning-based perception tasks (Yuan et al., 2019), and computer vision for safety (Leone et al., 2022). Technologies such as YOLO help improve ADAS by supporting pedestrian prediction and automatic braking for collision avoidance.

Redmon et al. (2016) explain that YOLO (You Only Look Once) is a real-time object detection system that treats detection as a single regression task within a convolutional neural network (CNN). Its grid-based design allows it to process full images at once, predicting bounding boxes and class probabilities efficiently, as illustrated in Figure 1. This makes it faster and more generalisable than models like Fast R-CNN (Ayachi et al., 2025; Redmon et al., 2016; Sapkota et al., 2025). The YOLOv5 nano version is lightweight and runs well on low-power embedded systems (Sapkota et al., 2025).

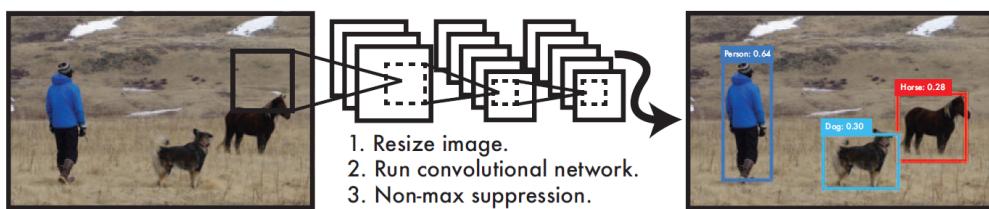


Figure 1: The YOLO Detection System (Redmon et al., 2016)

Object detection combines object classification and object localisation (Redmon et al., 2016). The system, based on CNNs trained on annotated datasets, outputs bounding boxes, class labels, and a confidence score between 0 and 1, indicating

detection certainty (Akshayaa & Nithin, 2021; Ayachi et al., 2025; Sazid & Afsha, 2023), as shown in Figure 1. Final detections depend on applying a confidence threshold to minimise false positives. Performance is typically assessed using mean Average Precision (mAP), recall and precision (Ayachi et al., 2025).

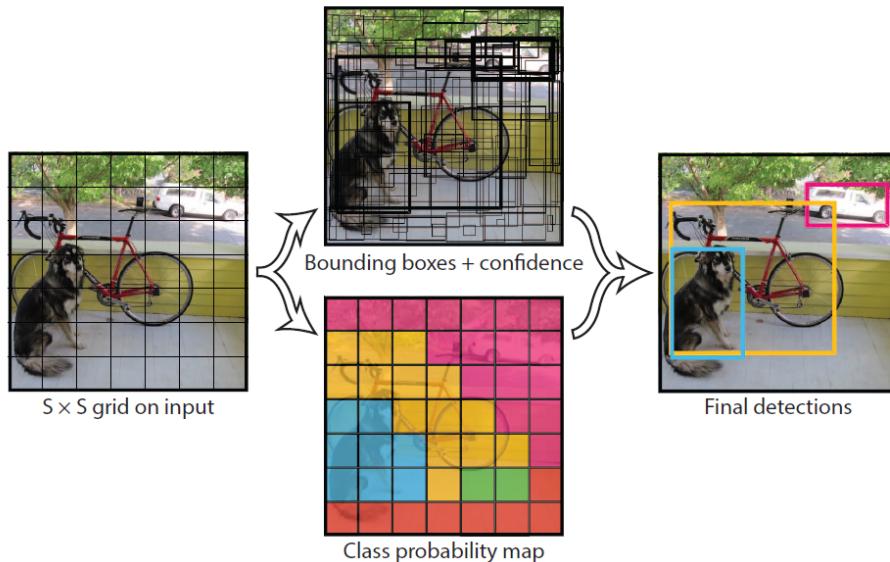


Figure 2: Bounding box prediction and confidence (Redmon et al., 2016)

3.2 Related Work

The following presents related works identified during the literature review.

Leone et al. (2022) demonstrated the feasibility of real-time safety monitoring using edge devices, but noted that their system still requires real-world validation.

Sazid and Afsha (2023) reported around 90% accuracy in controlled environments for their ADAS, though deployment on actual vehicle hardware has not yet been tested.

Lakshmi et al. (2023) identified significant challenges in real-life pedestrian

detection, including occlusion, lighting variation, and latency, despite strong performance in controlled conditions.

Sharma et al. (2022) emphasized the importance of detecting vulnerable pedestrians, particularly children, the elderly, and those with mobility issues, who are at higher risk of accidents.

Xu and Liu (2025) proposed an optimized YOLOv3 variant using G-modules and pruning to improve inference speed and compression, making it suitable for embedded applications, although it has not yet been tested specifically for pedestrian detection.

Reddy and Kumar (2024) developed a system using traditional image processing and Cascade Random Forest classification that performed well under specific lighting but lacked robustness for complex scenarios, illustrating an alternative approach to pedestrian detection beyond deep learning.

Miani et al. (2022) also discusses the modelling of assisted braking operations, which is one of the aspects this project aims to explore further.

3.3 Limitations and Gaps in the Literature

Across the literature, several gaps and limitations persist:

- **Limited evaluation** of systems in **real-world conditions**, including poor lighting, adverse weather, and crowded urban environments (Akshayaa & Nithin, 2021; Leone et al., 2022; Redmon et al., 2016; Sazid & Afsha, 2023).

- **Insufficient testing** on **embedded hardware** for real-time performance and scalability (Akshayaa & Nithin, 2021; Lakshmi et al., 2023; Sazid & Afsha, 2023).
- **Difficulty detecting small**, occluded or rapidly moving **objects** in cluttered or dynamic scenes (Lakshmi et al., 2023; Redmon et al., 2016).
- **High hardware cost** and **complexity** may limit adoption of advanced systems in consumer vehicles (Leone et al., 2022).
- **Lack of standardised** evaluation **protocols** across studies, making comparisons challenging (Lakshmi et al., 2023).
- **Limited validation** of pedestrian detection in **residential scenarios** such as driveways, where small or partially visible objects like children and pets present unique challenges (Reddy & Kumar, 2024; Redmon et al., 2016; Xu & Liu, 2025).

In summary, ITS leverage computer vision and deep learning to improve road safety, with real-time pedestrian detection being essential. Lightweight models like YOLOv5 nano, tailored for specific groups, enable effective use on embedded systems to enhance safety and reduce traffic accidents.

4 Methodology

4.1 Approach

Computer vision is widely used to interpret images and video, with object detection identifying and labelling items such as people or animals using bounding boxes. In this study, we applied it to detect children in vehicle blind spots, aiming to improve safety during driveway manoeuvres. We used Ultralytics' pre-trained YOLOv5n model (Jocher, 2020), which recognises 80 classes. While it detected older children as "person", it struggled with babies and toddlers, crucial for our intended use.

To address this, we re-trained the model by keeping the original classes and adding a new class, kids, totaling 81 classes as recommended by (Yasin, 2024). A new dataset was created by combining existing and new class images from (Roboflow, 2025), organised into train, validation, and test sets. The data.yaml file was updated with the correct number of classes and names, while the test set was kept separate to avoid data leakage.

Fine-tuning was performed starting from pre-trained COCO weights, with the option to freeze varying numbers of network layers (0 to 3) to protect the original classes. Training ran on a standard laptop (Intel® Core™ i5-1334U, 16 GB RAM, Windows 64-bit), with batch size set to 16 and 50 epochs and patience (early stop) 6. Freezing parts of the network helped maintain recognition of existing classes while enabling focused learning on the new kids class, balancing accuracy and processing

time.

The re-trained model (best.pt), developed through transfer learning and model fine-tuning, improved detection of babies and toddlers, achieving our main objective while still identifying older children as person in some cases, which is acceptable for this context. Animal detection capabilities were also retained. The final stage involved comparing model performance, preparing comparison tables, and refining the code for presentation. Despite some technical challenges, the solution remains lightweight and suitable for deployment on embedded systems.

The next step is integrating the model into passenger vehicles by linking it to the rear camera, multimedia system, alerts, and braking. Local processing needs compatible hardware like the low-cost Raspberry Pi or the more powerful NVIDIA Jetson Nano. According to Jain and Shah ([2021](#)), these platforms cost around 75 USD and 100 USD, respectively.

To prepare the Raspberry Pi for AI, for example, it is necessary install and update the 64-bit Raspberry Pi OS, enable the camera interface, and install Python 3 with pip along with essential AI libraries such as NumPy, SciPy, TensorFlow Lite, PyTorch (ARM versions), and OpenCV (Jain & Shah, [2021](#)).

4.2 System Design and Architecture

The proposed system architecture integrates computer vision technology with embedded hardware to create a comprehensive safety solution for driveway

environments. The design emphasises real-time processing capabilities whilst maintaining cost-effectiveness and ease of integration with existing vehicle systems.

Figure 3 illustrates the pipeline for YOLO model fine-tuning, demonstrating the systematic approach to enhancing detection capabilities through transfer learning and dataset augmentation.

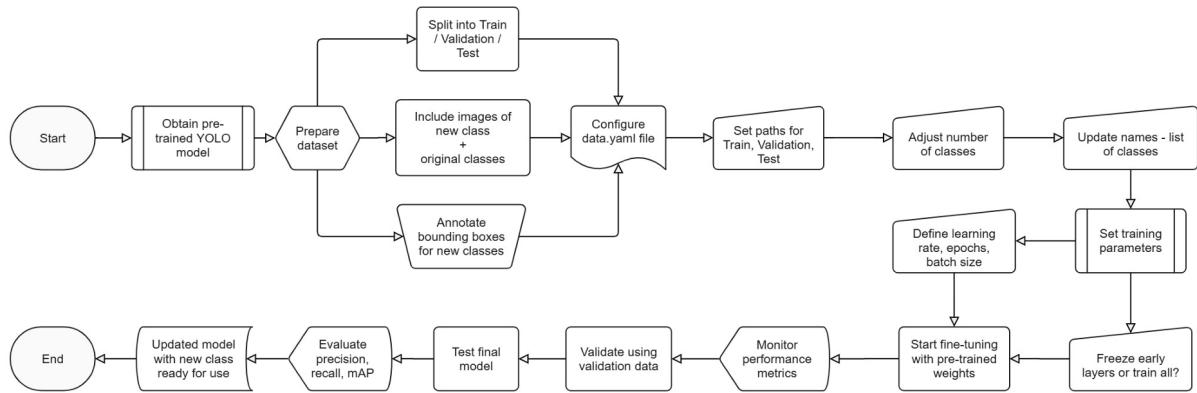


Figure 3: Pipeline for YOLO Model Fine-Tuning by Authors

Figure 4 presents the real-time object detection system architecture, showcasing the integration between camera input, processing unit, and alert mechanisms within the vehicle environment.

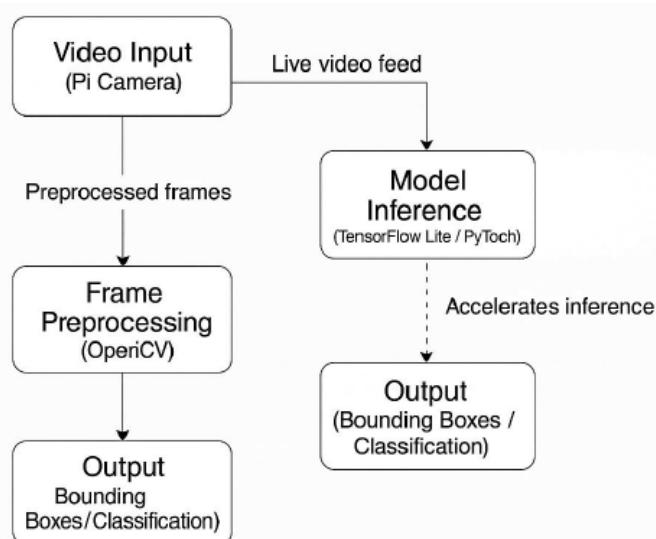


Figure 4: System Architecture for Real-Time Object Detection (Jain & Shah, 2021)

Figure 5 depicts the complete deployment system architecture, highlighting the interconnected components including embedded hardware, software modules, and vehicle integration points for practical implementation.

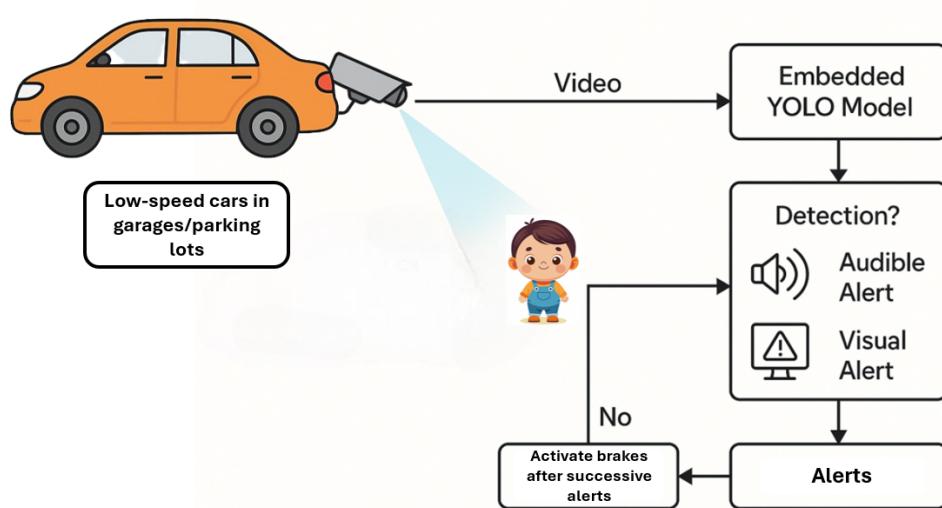


Figure 5: Architecture for Deployment system

4.3 Implementation Steps

1. Model Selection and Preparation

- Select YOLOv5n model from Ultralytics as base architecture.
- Evaluate pre-trained capabilities on 80 existing classes.
- Identify limitations in detecting babies and toddlers.

2. Dataset Enhancement

- Create a new "kids" class to supplement the existing "person" class.
- Combine the original dataset with new child-specific images.
- Organise data into train, validation, and test folders.

- Configure `data.yaml` file with updated class count (81 classes).

3. Model Fine-tuning

- Implement transfer learning from pre-trained COCO weights.
- Apply layer freezing (0–3 layers) to preserve original class recognition.
- Configure training parameters: batch size 16, 50 epochs.
- Execute training on standard hardware (Intel i5, 16 GB RAM).

4. Hardware Integration Planning

- Choose an embedded platform (Raspberry Pi or NVIDIA Jetson Nano).
- Install a compatible operating system.
- Set up a Python environment with essential AI libraries.

5. System Architecture Development

- Integrate rear camera with real-time video processing.
- Implement frame analysis pipeline for object detection.
- Configure alert mechanisms: visual, audible, vibrations, and automatic braking.
- Establish local processing workflow following IoT principles.

6. Performance Optimisation

- Adjust consecutive frame analysis settings.
- Fine-tune confidence threshold parameters.
- Balance detection sensitivity with false positive reduction.

- Test system responsiveness under various conditions.

Table 1 shows the layers for final solution from the user's perspective.

Table 1: Layers of the Driveway Safety Solution

Layer	Components
Hardware	<ul style="list-style-type: none"> • Raspberry Pi / NVIDIA Jetson Nano • Rear vehicle camera • Car control panel • Automatic braking system
Software/Model	<ul style="list-style-type: none"> • Fine-tuned YOLOv5n (best.pt) • OpenCV (Python) • PyTorch / TensorFlow Lite • Python libraries (NumPy, SciPy)
Integration	<ul style="list-style-type: none"> • Vehicle camera interface • Communication with vehicle electronic systems • Vehicle multimedia system • Braking control API
User Interface	<ul style="list-style-type: none"> • Visual alerts (GUI) • Audio notifications • Sensitivity settings • Confidence indicator

5 Results and Analysis

5.1 Software Features

The software developed combines computer vision and machine learning to enhance driveway safety. Using YOLOv5, it detects people, children, and pets in real time from the rear camera and alerts the driver to reduce accidents caused by limited visibility.

The system allows users to customise frame analysis and confidence thresholds to balance timely alerts and false positives. Alerts are visual and audible, and if ignored, the system can apply brakes at low speeds. It runs on a Raspberry Pi integrated with the vehicle's systems, with potential future improvements including haptic feedback and additional sensors.

5.2 Performance Metrics

Figure 6 illustrates an example image analysed by YOLOv5 before and after fine-tuning, highlighting improvements in detecting small or partially hidden children.

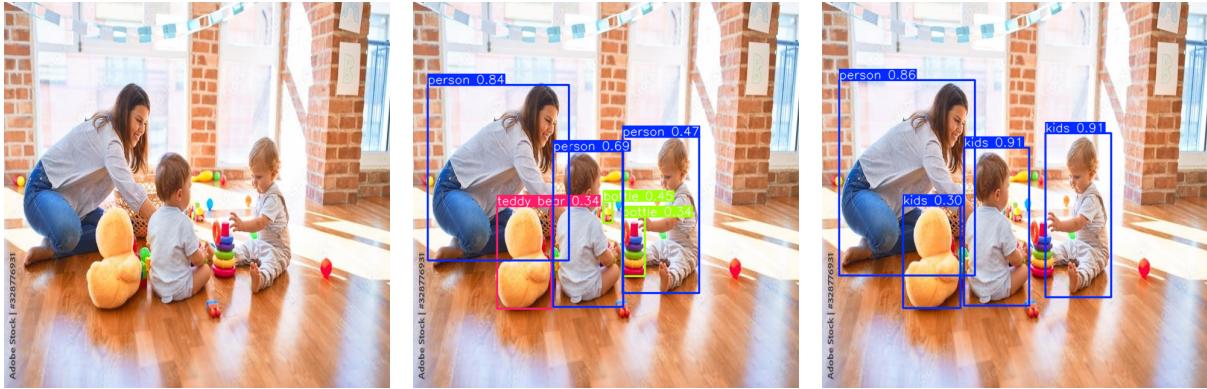


Figure 6: Original figure analysed by YOLOv5 before and after fine-tuning

The system's detection performance was assessed using a series of diagnostic plots, including the Precision-Recall curve (Figure 7), F1-Confidence curve (Figure 8), Recall-Confidence curve (Figure 9), and Precision-Confidence curve (Figure 10). Visual tests assessed the model's detection of children and pets at various confidence thresholds. Thresholds between 0.4 and 0.6 offered the best balance of accuracy and false positives, with 0.5 recommended. The model performed well in standard conditions, with low processing delay thanks to its lightweight design and on-device processing.

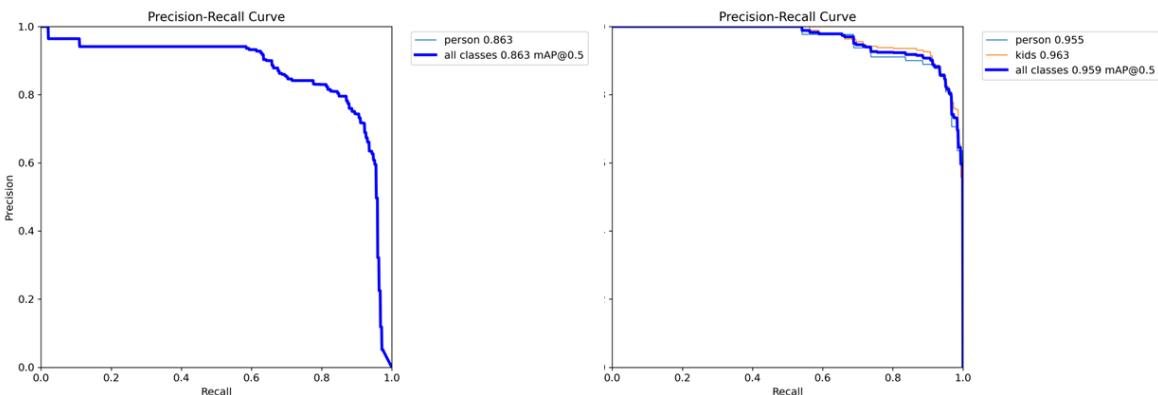


Figure 7: Precision-Recall Curve by authors using Python

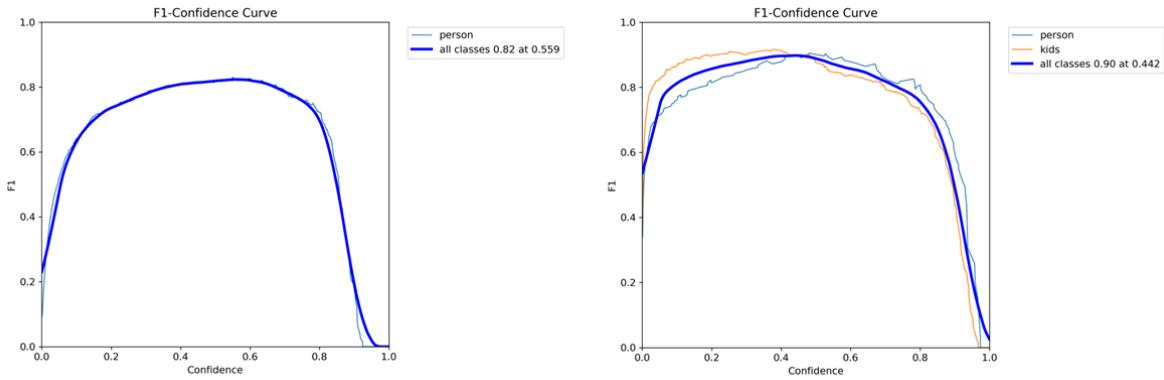


Figure 8: F1-Confidence Curve by authors using Python

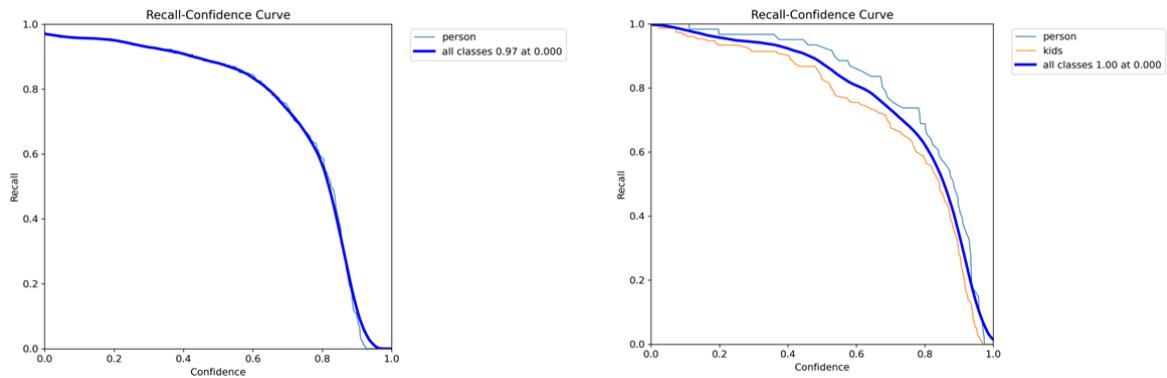


Figure 9: Recall-Confidence Curve by authors using Python

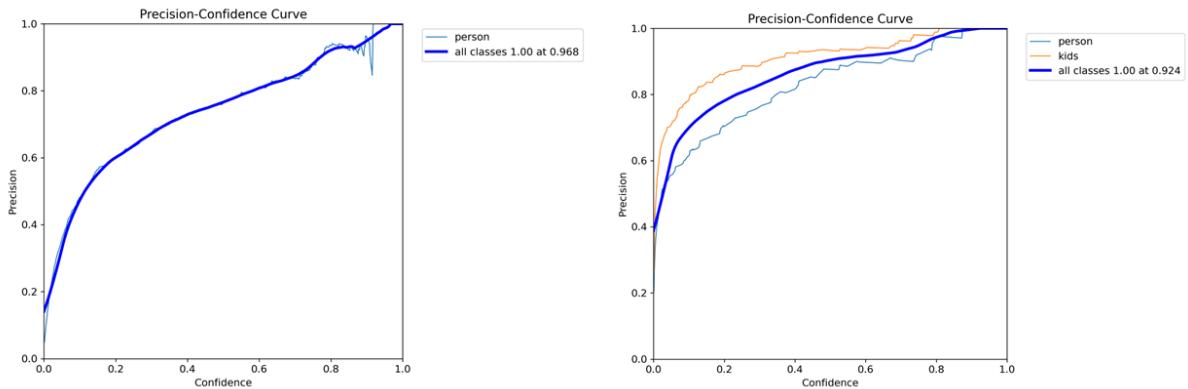


Figure 10: Precision-Confidence Curve by authors using Python

Latency stayed within real-time limits, consistently analysing rear-camera footage without lag. However, accuracy dropped in low light or cluttered scenes, especially with small or partially hidden children. This highlights the need for more varied training data

and suggests adding thermal sensors to improve detection in poor visibility conditions.

Overall, the model improved significantly after retraining, showing marked enhancements across almost all metrics and demonstrating strong potential for realistic applications. However, the system will require further refinement to ensure reliable performance across diverse conditions, as discussed in Section 6.2.

5.3 Security Assessment

Security was a key design focus, with all video processing done locally to protect privacy and comply with regulations like GDPR. No data is sent externally, reducing breach risks and cyber threats, though physical security remains important. Future work may add encryption and tamper detection to improve hardware protection.

Overall, the assessment shows strong privacy safeguards but highlights the need for enhanced hardware security through combined software and physical measures to ensure reliability in real use.

6 Discussion

6.1 Contributions and Challenges

These gaps inform the proposed AI system's development, which focuses on reliable, low-latency detection of vulnerable users in driveways using tailored computer vision and machine learning. Despite improvements, some limitations remain in the current project.

- **Class confusion:** The model occasionally confuses the “person” and “kids” classes, indicating for further refinement to better distinguish between these categories.
- **Small and occluded object detection:** Detecting very small or partially hidden children remains challenging, which may affect reliability in some scenes.
- **Limited training diversity:** The current training data may lack sufficient variety in lighting conditions and backgrounds, which could reduce the capacity to generalise to real-world scenarios.

Low-light performance: Thermal camera images can enhance detection in dark or obscured conditions by capturing heat signatures, allowing detection even in complete darkness, smoke, or fog, improving overall system robustness.

6.2 Future Enhancements

Future work should focus on addressing these limitations by expanding the dataset with more diverse and challenging samples, especially involving occluded children. Exploring more advanced model architectures and integrating multi-scale features could improve class differentiation. It is also important to optimise the model for real-time use on limited hardware while maintaining accuracy.

Incorporating temporal information from video and combining thermal with night vision cameras can improve detection stability and reduce misclassification. Night vision enhances low ambient light but struggles in total darkness, while thermal cameras offer complementary heat-based data. Together, these improvements can greatly boost system reliability and safety in complex real-world conditions.

7 Conclusion and Reflection

This project successfully developed an AI-driven driveway safety system using a fine-tuned YOLOv5n model to detect vulnerable road users, such as children and pets, in residential driveways. By enhancing detection capabilities through transfer learning and dataset augmentation, the system delivers real-time visual and audible alerts, with potential for automatic braking, addressing a critical gap in existing vehicle safety technologies.

Ethically, the system prioritizes user privacy by processing all video locally, ensuring compliance with GDPR and minimizing cybersecurity risks. The driver remains responsible for safe operation, reinforcing the system's role as an assistive tool.

Future work should focus on improving detection in challenging conditions and integrating additional sensors to enhance reliability. This project contributes to safer residential environments by reducing the risk of driveway accidents, particularly for vulnerable populations.

Individual Reflection

- **Alfredo Jr. Albis Torreña:** In this project, I contributed to the main idea of using AI/ML in car safety systems to reduce child deaths on driveways. While researching ITS projects, I found alarming data about this safety gap, despite modern car features. A groupmate shared a near-miss experience, which highlighted the issue's seriousness. I reviewed the proposed computer vision

model, assessing its strengths and weaknesses, and trained it on datasets I labelled myself. This showed me the model's strong potential to meet our goals and help prevent child accidents on driveways.

- **Eunseok Choi:** Throughout the project, I contributed mainly to the literature review, model selection, and dataset design. I selected and analysed relevant academic papers, organised the background into structured sections, and helped define the "kids" class for retraining. This process deepened my understanding of how object detection models work and how they can be adapted for real-world safety applications. I also gained experience in connecting research with design decisions and writing in an academic style.
- **Fabricio Mulato:** Based on gaps identified by colleagues, I contributed to selecting and retraining a pre-trained model to better detect children, using new images. I developed a script to test it on unseen data, making adjustments after colleagues tested it on different platforms and environments. I co-wrote the methodology and results sections and reviewed the report. I learned to retrain models, prepare datasets, and evaluate performance using appropriate metrics.
- **Yan Huang:** I contributed to the conclusion in the document. I collaborated on the methodology, evaluated the results, understood the code, reviewed the report, fixed the errors, and reduced redundancy. I learned how to develop with computer vision and how to train the model. The most important thing was learning how to collaborate with team members.

References

- Akshayaa, S., & Nithin, S. (2021). Comparative study of pedestrian detection techniques for driver assistance system. *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 1450–1454. <https://doi.org/10.1109/ICESC51422.2021.9532716>
- Ayachi, R., Said, Y., Afif, M., Alshammary, A., Hleili, M., & Abdelali, A. B. (2025). Assessing yolo models for real-time object detection in urban environments for advanced driver-assistance systems (adas). *Alexandria Engineering Journal*, 123, 530–549. <https://doi.org/10.1016/j.aej.2025.03.077>
- Brandao, M. (2019). Age and gender bias in pedestrian detection algorithms. <https://arxiv.org/abs/1906.10490>
- Jain, M., & Shah, A. (2021). Convolutional neural networks for real-time object detection with raspberry pi [International, Peer reviewed, Referred, Open access, ISSN Approved Journal]. *World Journal of Advanced Engineering Technology and Sciences*, 4(1), 87–105. <https://doi.org/10.30574/wjaets.2021.4.1.0067>
- Jocher, G. (2020). *Ultralytics yolov5* (Version 7.0) [Accessed: 24 June 2025]. <https://doi.org/10.5281/zenodo.3908559>
- KidsAndCars.org. (2023). *Facts - backovers* [Accessed: August 30, 2023]. <https://www.kidsandcars.org/backovers/facts>
- Lakshmi, D. S., Divya, A., Sreedevi, E., Kolagani, R., Gottumukkala, P., & Muddana, A. (2023). A study and analysis on pedestrian detection and tracking through

rear-view images. *Ingénierie des Systèmes d'Information*, 28(1), 23–30. <https://doi.org/10.18280/isi.280103>

Leone, G. R., Carboni, A., Nardi, S., & Moroni, D. (2022). Toward pervasive computer vision for intelligent transport system. *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, 26–29. <https://doi.org/10.1109/PerComWorkshops53856.2022.9767376>

Miani, M., Dunnhofer, M., Micheloni, C., Marini, A., & Baldo, N. (2022). Young drivers' pedestrian anti-collision braking operation data modelling for adas development. *Transportation Research Procedia*, 60, 432–439. <https://doi.org/10.1016/j.trpro.2021.12.056>

Otago Daily Times Online News. (2025). Road toll rises as toddler killed in driveway [Accessed: Jun. 30, 2025].

Perallos, A., Hernandez-Jayo, U., Onieva, E., & García-Zuazola, I. J. (Eds.). (2016). *Intelligent transport systems: Technologies and applications*. John Wiley; Sons, Ltd.

Reddy, P., & Kumar, A. (2024). Pedestrian detection system using image processing and cascade random forest. *International Journal of Intelligent Systems*, 39(2), 215–223.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. <https://doi.org/10.1109/CVPR.2016.91>

Roboflow. (2025). Roboflow: Computer vision tools for developers and enterprises [Accessed: 25 June 2025]. <https://roboflow.com/>

Sapkota, R., et al. (2025). Yolo advances to its genesis: A decadal and comprehensive review of the you only look once (yolo) series. *Artificial Intelligence Review*, 58(9), 274. <https://doi.org/10.1007/s10462-025-11253-3>

Sazid, M. M., & Afsha, S. (2023). Advanced driving assistance system using computer vision and deep learning algorithms. *2023 Third International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)*, 17–23. <https://doi.org/10.1109/ICUIS60567.2023.00011>

Sharma, D., Hade, T., & Tian, Q. (2022). Comparison of deep object detectors on a new vulnerable pedestrian dataset. *arXiv preprint*. <https://doi.org/10.48550/ARXIV.2212.06218>

Tahir, G. A. (2024). Ethical challenges in computer vision: Ensuring privacy and mitigating bias in publicly available datasets. <https://arxiv.org/abs/2409.10533>

Xu, C., & Liu, Y. (2025). Lightweight yolov3 for embedded systems using g-modules and channel pruning. *Proceedings of the 2025 International Conference on Embedded AI Systems*, 100–108.

Yasin, M. (2024, March 4). *Extending yolo models with new classes through additional head* [Accessed: 23 June 2025]. <https://doi.org/10.13140/RG.2.2.12569.53609>

Yuan, T., da Rocha Neto, W. B., Rothenberg, C., Obraczka, K., Barakat, C., et al. (2019). Machine learning for next-generation intelligent transportation systems: A survey

[Available at HAL: <https://hal.archives-ouvertes.fr/hal-02284820v1>]. *arXiv preprint.*

Appendix A: Training Results and Visualisations

Training ended at epoch 45 due to early stopping with a patience of 6, as no improvement occurred during that period.

epoch	metrics/precision(B)	metrics/recall(B)	metrics/mAP50(B)	metrics/mAP50-95(B)
1	0	0	0	0
2	0	0	0	0
3	0.32707	0.97856	0.56954	0.38867
4	0.49009	0.88459	0.60523	0.41209
5	0.53266	0.77654	0.62517	0.4286
6	0.49837	0.88685	0.61666	0.41152
7	0.61449	0.73342	0.7273	0.47477
8	0.61877	0.75312	0.75413	0.49984
9	0.69548	0.72956	0.78785	0.52516
10	0.73054	0.75347	0.84946	0.56041
11	0.75943	0.81736	0.87098	0.56612
12	0.74591	0.84355	0.88353	0.58417
13	0.82451	0.81869	0.89427	0.56009
14	0.79288	0.86276	0.90775	0.60537
15	0.85209	0.82661	0.91171	0.59563
16	0.84611	0.8366	0.90681	0.60876
17	0.84874	0.84636	0.91865	0.61169
18	0.70658	0.86771	0.89338	0.61485
19	0.85118	0.87656	0.92216	0.62261
20	0.8458	0.90022	0.93506	0.63969
21	0.85005	0.81473	0.92372	0.63761
22	0.8183	0.90122	0.9315	0.63055
23	0.8325	0.92031	0.94458	0.65345
24	0.86604	0.91493	0.9484	0.65978
25	0.8473	0.866	0.93766	0.66631
26	0.86869	0.85239	0.94472	0.63895
27	0.84441	0.81676	0.92632	0.65499
28	0.82102	0.89254	0.94346	0.67504
29	0.90955	0.88199	0.95527	0.66082
30	0.80397	0.92664	0.94037	0.67694
31	0.88526	0.92243	0.96157	0.68913
32	0.84966	0.91896	0.95773	0.68047
33	0.83216	0.88065	0.93958	0.67493
34	0.8475	0.92562	0.95751	0.68392
35	0.91514	0.81946	0.94954	0.68282
36	0.89049	0.91076	0.9589	0.69139
37	0.87659	0.92392	0.95791	0.68736
38	0.86504	0.90083	0.95427	0.6859
39	0.89128	0.90864	0.95884	0.69194
40	0.90098	0.90458	0.96117	0.68751
41	0.82441	0.92005	0.95765	0.68525
42	0.7882	0.94467	0.95057	0.67524
43	0.92697	0.85023	0.96073	0.67276
44	0.90548	0.88318	0.96113	0.6714
45	0.90927	0.89104	0.95787	0.67044

Table 2: Metrics after each training epoch.



Figure 11: Test batch for the original YOLOv5n.pt model



Figure 12: Test batch for the fine-tuned best.pt model

Appendix B: Complete Project Source Code

01_load_test_orig.ipynb

Objective

This notebook aims to load the pre-trained YOLOv5n model (`yolov5n.pt`) already downloaded locally and test its inference on sample images and videos without any fine-tuning.

The YOLOv5n pre-trained model (80 classes) was downloaded from Ultralytics. For this stage of the project, only the classes of interest were selected: `['person', 'cat', 'dog', 'horse', 'sheep', 'cow']`.

Steps:

1. Import the necessary libraries
2. Load the pre-trained YOLOv5n model
3. Test the model on sample images and videos
4. Display the results

```
In [5]: # 1. Import required libraries
import torch
from pathlib import Path
import cv2
from matplotlib import pyplot as plt

In [6]: # 2. Load the pre-trained YOLOv5n model
model_path = Path("./yolov5n.pt")
assert model_path.is_file(), "Model not found at {}"

# Load the model
model = torch.hub.load('ultralytics/yolov5', 'custom', path=str(model_path), force_reload=False)

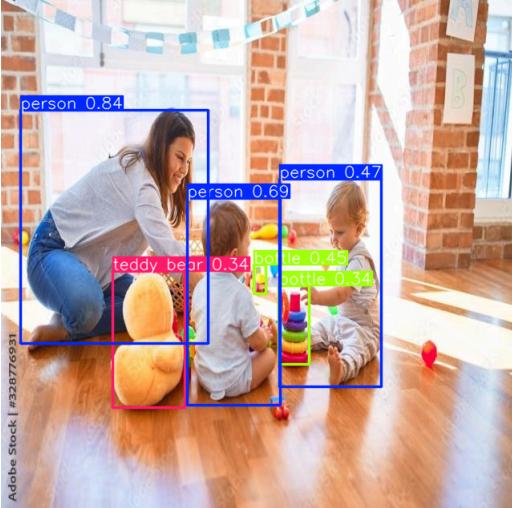
Using cache found in C:\Users\fmula\.cache\torch\hub\ultralytics_yolov5_master
C:\Users\fmula\.cache\torch\hub\ultralytics_yolov5_master\utils\general.py:32: UserWarning: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources package is slated for removal as early as 2025-11-30. Refrain from using this package or pin to Setuptools<81.
import pkg_resources as pkg
YOLOv5 2025-6-29 Python-3.13.5 torch-2.7.1+cpu CPU

Fusing layers...
YOLOv5n summary: 213 layers, 1867405 parameters, 0 gradients, 4.5 GFLOPs
Adding AutoShape...

In [7]: # 3. Test the model on a sample image
# For this example, place an image named "test_image.jpg" in the same folder
img_path = Path("./img1.jpg")
assert img_path.is_file(), "Test image not found at {}"

# Run inference
results = model(str(img_path))

C:\Users\fmula\.cache\torch\hub\ultralytics_yolov5_master\models\common.py:906: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
with amp.autocast(autocast):
```



Adobe Stock | #328776931

```
In [6]: # You can also access detailed results (e.g., labels, bounding boxes)
print(results.pandas().xyxy[0]) # pandas DataFrame with bounding boxes and confidence scores
```

	xmin	ymin	xmax	ymax	confidence	class
0	195.552307	435.718018	295.22137	713.602417	0.901315	0
1	23.848667	426.37107	141.853851	698.408020	0.895987	0
2	182.293594	128.703696	298.528625	413.254395	0.894124	0
3	853.200500	8.717178	1032.884033	384.004089	0.883776	0
4	120.854164	519.246765	208.705597	708.523743	0.880960	0
5	663.791382	341.769653	871.459598	687.699497	0.844776	0
6	532.495728	499.381958	661.826412	713.919490	0.837192	0
7	1817.397095	467.970520	1131.847720	680.988020	0.785037	0
8	1144.698126	507.720459	1229.681685	694.389356	0.765824	0
9	982.253967	468.119169	1832.966675	521.736816	0.274979	29
10	1866.734253	511.590982	1128.867310	680.152344	0.266940	0

name
0 person
1 person
2 person
3 person
4 person
5 person
6 person
7 person
8 person
9 frisbee
10 person

```
In [ ]: # Showing some metrics for original model
from ultralytics import YOLO

model = YOLO('yolov5n.pt')
metrics = model.val(data='dataset/data_original.yaml', task='test', batch=16)
print(metrics)
```

PRO TIP: Replace `'model=yolov5n.pt'` with new `'model=yolov5n.pt'`.
YOLOv5 'u' models are trained with <https://github.com/ultralytics/yolov5>.

```
Ultralytics 8.3.160 Python-3.13.5 torch-2.7.1+cpu CPU (13th Gen Intel Core(TM) i5-1334U)
YOLOv5n summary (fused): 84 layers, 2,649,200 parameters, 0 gradients, 7.7 GFLOPs
val: Fast image access (ping: 0.000 ms, read: 72.231.7 MB/s, size: 42 KB)
val: Scanning C:\Users\fmula\Documents\YooobeSubjects\Term_2\Intelligent_TS\Assessment_2\ITS2_AI_Driveway\dataset\test\labels... 117 images, 2 backgrounds, 0 corrupt: 100%|██████████| 117/117 [00:00<00:00, 1849.63it/s]
val: New cache created: C:\Users\fmula\Documents\YooobeSubjects\Term_2\Intelligent_TS\Assessment_2\ITS2_AI_Driveway\dataset\test\labels.cache
```

Class	Images	Instances	Box(P)	R	mAP50	mAP50-95
	8/8	[00:08<00:00, 1.00s/it]				

```

      all    117    246    0.792    0.854    0.863    0.5
    person   115    246    0.792    0.854    0.863    0.5
Speed: 1.1ms preprocess, 58.2ms inference, 0.0ms loss, 2.3ms postprocess per image

```

Speed: 1.1ms preprocess, 58.2ms inference, 0.0ms loss, 2.3ms postprocess per image
Results saved to runs\test\val5

ultralytics.utils.metrics.DetMetrics object with attributes:

www.123RF.com

2494,	0.82525,	0.82552,	0.82577,	0.82602,	0.82627,	0.82383,	0.82321,	0.8226,	0.8215,	0.82089,	0.8214,	0.82048,	0.81955,	0.81908,	0.81879,	0.8185,	0.81822,	0.8
1793,	0.81764,	0.81736,	0.81707,	0.81732,	0.81834,	0.81847,	0.81863,	0.81643,	0.81683,	0.81784,	0.81941,	0.8192,	0.81853,	0.81918,	0.82056,	0.82012,	0.81857,	0.81708,
1377,	0.81269,	0.81163,	0.81181,	0.81193,	0.81225,	0.81235,	0.81213,	0.8102,	0.81044,	0.81005,	0.80967,	0.80928,	0.8089,	0.80851,	0.80839,	0.79971,	0.79933,	0.79896,
0483,	0.80272,	0.80302,	0.80332,	0.80362,	0.80893,	0.80852,	0.80859,	0.80859,	0.80853,	0.80822,	0.80839,	0.79971,	0.79933,	0.79896,	0.80399,	0.808372,	0.80839,	0.80427,
9746,	0.80843,	0.808472,	0.80852,	0.80859,	0.80859,	0.808731,	0.808709,	0.8065,	0.80591,	0.80533,	0.8022,	0.80039,	0.79971,	0.79933,	0.79896,	0.79859,	0.79821,	0.79784,
8283,	0.77551,	0.77231,	0.77668,	0.7716,	0.77319,	0.77668,	0.77102,	0.77102,	0.77179,	0.7712,	0.77062,	0.77004,	0.778633,	0.778491,	0.77852,	0.78542,	0.78548,	0.78407,
6882,	0.76838,	0.76787,	0.76666,	0.76431,	0.76342,	0.76056,	0.75922,	0.76022,	0.76235,	0.76145,	0.76056,	0.75936,	0.76258,	0.76271,	0.76194,	0.76644,	0.75976,	0.75913,
7585,	0.75787,	0.75707,	0.75606,	0.75504,	0.75718,	0.75606,	0.75597,	0.75606,	0.75816,	0.75649,	0.75165,	0.75055,	0.74945,	0.7505,	0.74951,	0.74852,	0.74756,	0.74694,
4508,	0.75641,	0.75749,	0.75598,	0.75732,	0.75962,	0.75928,	0.75951,	0.759487,	0.759487,	0.75926,	0.759179,	0.75912,	0.759062,	0.759004,	0.758633,	0.758491,	0.75852,	0.75847,
7219,	0.72125,	0.7206,	0.72016,	0.72119,	0.71436,	0.71716,	0.71601,	0.71601,	0.71606,	0.716961,	0.71684,	0.716863,	0.716837,	0.716834,	0.716849,	0.716892,	0.716934,	0.716766,
6541,	0.66357,	0.6557,	0.6524,	0.65124,	0.65090,	0.65351,	0.64371,	0.64195,	0.64351,	0.63176,	0.62641,	0.61793,	0.61879,	0.59947,	0.59343,	0.58785,	0.58265,	0.57819,
6565,	0.56167,	0.54986,	0.54719,	0.54534,	0.54015,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,	0.53503,
8639,	0.38235,	0.37479,	0.37151,	0.36359,	0.35381,	0.32843,	0.31099,	0.31737,	0.31168,	0.30503,	0.29862,	0.29287,	0.2835,	0.2616,	0.24562,	0.23769,	0.21967,	0.21248,
0358,	0.20263,	0.20898,	0.19773,	0.19665,	0.19654,	0.18681,	0.17818,	0.169473,	0.15552,	0.1423,	0.13577,	0.13238,	0.12898,	0.11786,	0.077742,	0.070936,	0.060275,	0.055922,
3881,	0.039738,	0.03982,	0.028622,	0.023361,	0.021952,	0.016296,	0.011713,	0.016296,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
0.0004004,	0.00050805,	0.0006006,	0.007007,	0.008008,	0.009009,	0.01001,	0.011011,	0.012012,	0.013013,	0.014014,	0.015015,	0.016016,	0.017017,	0.018018,	0.019019,	0.02002,	0.021021,	0.
020222,	0.023022,	0.024024,	0.025025,	0.026026,	0.027027,	0.028028,	0.029029,	0.03003,	0.031031,	0.032032,	0.033033,	0.034034,	0.035035,	0.036036,	0.037037,	0.038038,	0.039039,	0.04004,
041,	0.042042,	0.043043,	0.044044,	0.045045,	0.046046,	0.047047,	0.048048,	0.049049,	0.05005,	0.051051,	0.052052,	0.053053,	0.054054,	0.055055,	0.056056,	0.057057,	0.058058,	0.059059,
065,	0.065665,	0.065765,	0.065868,	0.066069,	0.067078,	0.067171,	0.067272,	0.067373,	0.067474,	0.067575,	0.067676,	0.067777,	0.067878,	0.067979,	0.068008,	0.0681081,	0.0682082,	0.0683083,
089,	0.09099,	0.091891,	0.092092,	0.093093,	0.094094,	0.095095,	0.096096,	0.097097,	0.098098,	0.099099,	0.1001,	0.1011,	0.1021,	0.1031,	0.1041,	0.10511,	0.10611,	0.10711,
311,	0.11411,	0.11512,	0.11612,	0.11712,	0.11812,	0.11912,	0.12012,	0.12112,	0.12212,	0.12312,	0.12412,	0.12513,	0.12613,	0.12713,	0.12813,	0.12913,	0.13013,	0.13113,
714,	0.13814,	0.13914,	0.14014,	0.14114,	0.14214,	0.14314,	0.14414,	0.14514,	0.14615,	0.14715,	0.14815,	0.14915,	0.15015,	0.15115,	0.15215,	0.15315,	0.15415,	0.15516,
116,	0.16216,	0.16316,	0.16416,	0.16517,	0.16617,	0.16717,	0.16817,	0.16917,	0.17017,	0.17117,	0.17217,	0.17317,	0.17417,	0.17518,	0.17618,	0.17718,	0.17818,	0.17918,
519,	0.18619,	0.18719,	0.18819,	0.18919,	0.19019,	0.19119,	0.19219,	0.19319,	0.19419,	0.19519,	0.19619,	0.19719,	0.19819,	0.19919,	0.20019,	0.20119,	0.20219,	0.20319,
921,	0.20201,	0.21121,	0.21221,	0.21321,	0.21421,	0.21521,	0.21621,	0.21721,	0.21821,	0.21921,	0.22021,	0.22121,	0.22221,	0.22321,	0.22421,	0.22521,	0.22621,	0.22721,
323,	0.23423,	0.23524,	0.23624,	0.23724,	0.23824,	0.23924,	0.24024,	0.24124,	0.24224,	0.24324,	0.24424,	0.24525,	0.24625,	0.24725,	0.24825,	0.24925,	0.25025,	0.25125,
726,	0.25826,	0.25926,	0.26026,	0.26126,	0.26226,	0.26326,	0.26426,	0.26526,	0.26626,	0.26726,	0.26826,	0.26926,	0.27026,	0.27127,	0.27227,	0.27327,	0.27427,	0.27528,
128,	0.28228,	0.28328,	0.28428,	0.28529,	0.28629,	0.28729,	0.28829,	0.28929,	0.29029,	0.29129,	0.29229,	0.29329,	0.29429,	0.29529,	0.29629,	0.29729,	0.29829,	0.29929,
531,	0.30631,	0.30731,	0.30831,	0.30931,	0.31031,	0.31131,	0.31231,	0.31331,	0.31431,	0.31532,	0.31632,	0.31732,	0.31832,	0.31932,	0.32032,	0.32132,	0.32232,	0.32332,
933,	0.33333,	0.33433,	0.33533,	0.33633,	0.33733,	0.33833,	0.33933,	0.34033,	0.34133,	0.34233,	0.34334,	0.34434,	0.34534,	0.34634,	0.34734,	0.34834,	0.34934,	0.35034,
335,	0.35453,	0.35556,	0.35656,	0.35756,	0.35856,	0.35956,	0.36056,	0.36156,	0.36256,	0.36356,	0.36456,	0.36556,	0.36656,	0.36756,	0.36856,	0.36956,	0.37056,	0.37156,
738,	0.37938,	0.37938,	0.38038,	0.38138,	0.38238,	0.38338,	0.38438,	0.38538,	0.38638,	0.38738,	0.38838,	0.38938,	0.39039,	0.39139,	0.39239,	0.39339,	0.39439,	0.3954,
014,	0.40424,	0.40434,	0.40444,	0.40454,	0.40464,	0.40474,	0.40484,	0.40494,	0.40504,	0.40514,	0.40524,	0.40534,	0.40544,	0.40554,	0.40564,	0.40574,	0.40584,	0.40594,
543,	0.42463,	0.42473,	0.42483,	0.42493,	0.42503,	0.42513,	0.42523,	0.42533,	0.42543,	0.42553,	0.42563,	0.42573,	0.42583,	0.42593,	0.42603,	0.42613,	0.42623,	0.42632,
945,	0.45045,	0.45145,	0.45245,	0.45345,	0.45445,	0.45545,	0.45645,	0.45745,	0.45845,	0.45945,	0.46045,	0.46146,	0.46246,	0.46346,	0.46446,	0.46546,	0.46646,	0.46745,
347,	0.47447,	0.47548,	0.47648,	0.47748,	0.47848,	0.47948,	0.48048,	0.48148,	0.48248,	0.48348,	0.48448,	0.48548,	0.48648,	0.48748,	0.48848,	0.48948,	0.49049,	0.49149,
975,	0.4985,	0.4995,	0.5005,	0.5015,	0.5025,	0.5035,	0.5045,	0.5055,	0.5065,	0.5075,	0.5085,	0.5095,	0.51051,	0.51151,	0.51251,	0.51351,	0.51451,	0.51552,
152,	0.52252,	0.52352,	0.52453,	0.52553,	0.52653,	0.52753,	0.52853,	0.52953,	0.53053,	0.53153,	0.53253,	0.53353,	0.53453,	0.53554,	0.53654,	0.53754,	0.53854,	0.53954,
555,	0.52853,	0.52953,	0.53053,	0.53153,	0.53253,	0.53353,	0.53453,	0.53553,	0.53653,	0.53753,	0.53853,	0.53953,	0.54053,	0.54153,	0.54254,	0.54354,	0.54454,	0.54554,
957,	0.57057,	0.57157,	0.57257,	0.57357,	0.57457,	0.57557,	0.57657,	0.57757,	0.57857,	0.57957,	0.58058,	0.58158,	0.58258,	0.58358,	0.58458,	0.58559,	0.58659,	0.58759,
359,	0.59459,	0.59556,	0.59656,	0.59756,	0.59856,	0.59956,	0.60056,	0.60156,	0.60256,	0.60356,	0.60456,	0.60556,	0.60656,	0.60756,	0.60856,	0.60956,	0.61056,	0.61156,
762,	0.61862,	0.62063,	0.62263,	0.62463,	0.62663,	0.62863,	0.63063,	0.63263,	0.63463,	0.63663,	0.63863,	0.64063,	0.64263,	0.64463,	0.64663,	0.64863,	0.65063,	0.65263,
164,	0.64264,	0.64364,	0.64464,	0.64565,	0.64665,	0.64765,	0.64865,	0.64965,	0.65065,	0.65165,	0.65265,	0.65365,	0.65465,	0.65565,	0.65666,	0.65766,	0.65866,	0.65966,
567,	0.66667,	0.66767,	0.66867,	0.66967,	0.67067,	0.67167,	0.67267,	0.67367,	0.67467,	0.67567,	0.67667,	0.67767,	0.67867,	0.67967,	0.68068,	0.68168,	0.68268,	0.68368,
969,	0.69689,	0.69719,	0.69749,	0.69779,	0.69809,	0.69819,	0.69829,	0.69839,	0.69849,	0.69859,	0.69869,	0.69879,	0.69889,	0.69899,	0.69909,	0.69919,		

02_finetune.ipynb

Objective

This notebook performs fine-tuning of the YOLOv5n model using a custom dataset.
The training configuration includes:

- Custom data YAML file
- 1 training epoch (for demonstration purposes)
- Batch size of 2
- Initial learning rate of 0.01
- Freeze layers: [0, 1, 2] (backbone layers frozen)
- Early stopping patience of 3

Steps:

1. Import necessary libraries
2. Load the YOLOv5 model
3. Train (fine-tune) the model
4. Save training results

```
In [2]: # 1. Import necessary Libraries
import torch
from pathlib import Path

In [1]: # # 2. Load the YOLOv5 model
# model_path = Path("./yolov5n.pt")
# assert model_path.is_file(), f"Model not found at {model_path}"
# model = torch.hub.load("ultralytics/yolov5", "custom", path=str(model_path), force_reload=True)

Downloading: "https://github.com/ultralytics/yolov5/zipball/master" to C:\Users\fmula/.cache\torch\hub\master.zip
C:\Users\fmula\.cache\torch\hub\ultralytics_yolov5_master\utils\general.py:32: UserWarning: pkg_resources is deprecated as an API. See https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources package is slated for removal as early as 2025-11-30. Refrain from using this package or pin to Setuptools<81.
import pkg_resources as pkg
YOLOv5 2025-6-29 Python-3.13.5 torch-2.7.1+cpu CPU

Fusing layers...
YOLOv5n summary: 213 layers, 1867405 parameters, 0 gradients, 4.5 GFLOPs
Adding AutoShape...

In [6]: # 2. Load the YOLOv5 model
from ultralytics import YOLO

model = YOLO('yolov5n.pt') # ou o caminho do modelo

PRO TIP Replace 'model=yolov5n.pt' with new 'model=yolov5nu.pt'.
YOLOv5 'u' models are trained with https://github.com/ultralytics/ultralytics and feature improved performance vs standard YOLOv5 models trained with https://github.com/ultralytics/yolov5.

Downloading https://github.com/ultralytics/assets/releases/download/v8.3.0/yolov5nu.pt to 'yolov5nu.pt'...
100%[██████████] 5.31M/5.31M [00:00<00:00, 16.8MB/s]

In [8]: # 3. Set training parameters

# Run training
results = model.train(
    data="dataset/data_fine_tuning.yaml",
    epochs=50,
    batch=16,
    lr=0.001,
    freeze=[0, 1, 2, 3], # Congelar camadas iniciais
    patience=6
)

Ultralytics 8.3.160 Python-3.13.5 torch-2.7.1+cpu CPU (13th Gen Intel Core(TM) i5-1334U)
engine/trainer agnostic_nms=False, amp=True, augment=False, auto_mosaic=False, batch_size=16, bgr=0.0, box=7.5, cache=False, cfg=None, classes=None, close_mosaic=10, cls=0.5, conf=None, copy_paste=0.0, copy_paste_mode=flip, cos_lr=False, cutmix=0.0, data=dataset/data_fine_tuning.yaml, degrees=0.0, deterministic=True, device=cpu, dfl=1.5, dnrm=False, dropout=0.0, dynamic=False, embed=None, epochs=50, erasing=0.4, exist_ok=False, flipl=0.5, filipd=0.0, format=torchscript, ignore_val=False, project=True, rect=False, resume=False, retina_masks=False, save=True, save_conf=False, save_crop=False, save_dir=runs\detect\train2, save_frames=False, save_json=False, save_period=-1, save_txt=False, scale=0.5, seed=0, shear=0.0, skip=0, show=False, show_boxes=True, show_conf=True, show_labels=True, simplify=True, single_cls=False, source=None, split=val, stream_buffer=False, task=detect, time=None, tracker=botsort.yaml, translate=0.1, val=True, verbose=True, vid_stride=1, visualize=False, warmup_bias_lr=0.1, warmup_epochs=3.0, warmup_momentum=0.8, weight_decay=0.0005, workers=8, workspace=None
Overriding model.yaml nc=88 with nc=81

          from n      params   module                                arguments
  0       -1 1        1760  ultralytics.nn.modules.conv.Conv     [3, 16, 6, 2, 2]
  1       -1 1        4672  ultralytics.nn.modules.conv.Conv     [16, 32, 3, 2]
  2       -1 1        4800  ultralytics.nn.modules.block.C3     [32, 32, 1]
  3       -1 1       18560  ultralytics.nn.modules.conv.Conv     [32, 64, 3, 2]
  4       -1 2        29184  ultralytics.nn.modules.block.C3     [64, 64, 2]
  5       -1 1       73984  ultralytics.nn.modules.conv.Conv     [64, 128, 3, 2]
  6       -1 3       156928  ultralytics.nn.modules.block.C3     [128, 128, 3]
  7       -1 1       295424  ultralytics.nn.modules.conv.Conv     [128, 256, 3, 2]
  8       -1 1       296448  ultralytics.nn.modules.block.C3     [256, 256, 1]
  9       -1 1       164608  ultralytics.nn.modules.block.SPPF    [256, 256, 5]
 10      -1 1       33824  ultralytics.nn.modules.conv.Conv     [256, 128, 1, 1]
 11      -1 1        0      torch.nn.modules.upsampling.Upsample [None, 2, 'nearest']
 12     [-1, 6] 1        0      ultralytics.nn.modules.Concat     [1]
 13      -1 1       908880  ultralytics.nn.modules.block.C3     [256, 128, 1, False]
 14      -1 1       8320  ultralytics.nn.modules.conv.Conv     [128, 64, 1, 1]
 15      -1 1        0      torch.nn.modules.upsampling.Upsample [None, 2, 'nearest']
 16     [-1, 4] 1        0      ultralytics.nn.modules.Concat     [1]
 17      -1 1       22912  ultralytics.nn.modules.block.C3     [128, 64, 1, False]
 18      -1 1       36932  ultralytics.nn.modules.conv.Conv     [64, 64, 3, 2]
 19     [-1, 14] 1        0      ultralytics.nn.modules.ConvCat    [1]
 20      -1 1       744096  ultralytics.nn.modules.block.C3     [128, 128, 1, False]
 21      -1 1       147712  ultralytics.nn.modules.conv.Conv     [128, 128, 3, 2]
 22     [-1, 10] 1        0      ultralytics.nn.modules.Concat    [1]
 23      -1 1       296448  ultralytics.nn.modules.block.C3     [256, 256, 1, False]
 24     [17, 20, 23] 1       906541  ultralytics.nn.modules.head.Detect [81, [64, 128, 256]]
YOLOv5n summary: 153 layers, 2,663,693 parameters, 2,663,677 gradients, 7.9 GFLOPs

Transferring 391/427 items from pretrained weights
Freezing layer 'model.0.conv.weight'
Freezing layer 'model.0.bn.weight'
Freezing layer 'model.0.bn.bias'
Freezing layer 'model.1.conv.weight'
Freezing layer 'model.1.bn.weight'
Freezing layer 'model.1.bn.bias'
Freezing layer 'model.2.cv1.conv.weight'
Freezing layer 'model.2.cv1.bn.weight'
Freezing layer 'model.2.cv1.bn.bias'
Freezing layer 'model.2.cv2.conv.weight'
Freezing layer 'model.2.cv2.bn.weight'
Freezing layer 'model.2.cv2.bn.bias'
Freezing layer 'model.2.cv3.conv.weight'
Freezing layer 'model.2.cv3.bn.weight'
Freezing layer 'model.2.cv3.bn.bias'
Freezing layer 'model.2.m.0.cv1.conv.weight'
Freezing layer 'model.2.m.0.cv1.bn.weight'
Freezing layer 'model.2.m.0.cv1.bn.bias'
Freezing layer 'model.2.m.0.cv2.conv.weight'
Freezing layer 'model.2.m.0.cv2.bn.weight'
Freezing layer 'model.2.m.0.cv2.bn.bias'
Freezing layer 'model.2.m.0.cv3.bn.weight'
Freezing layer 'model.2.m.0.bn.weight'
Freezing layer 'model.3.bn.bias'
Freezing layer 'model.24.dfl.conv.weight'
train: Fast image access (ping: 0.000 ms, read: 630.3225.4 MB/s, size: 53.7 KB)
train: Scanning C:\Users\fmula\Documents\Yoobee\Subjects\Term_2\Intelligent_TS\Assessment_2\ITS2_AI_Driveway\dataset\train_tuned\labels... 378 images, 0 backgrounds, 0 corrupt: 100%[██████████] 378/378 [00:00<00:00, 2129.32it/s]
train: New cache created: C:\Users\fmula\Documents\Yoobee\Subjects\Term_2\Intelligent_TS\Assessment_2\ITS2_AI_Driveway\dataset\train_tuned\labels.cache
val: Fast image access (ping: 0.382 ms, read: 76.418.4 MB/s, size: 52.5 KB)
val: Scanning C:\Users\fmula\Documents\Yoobee\Subjects\Term_2\Intelligent_TS\Assessment_2\ITS2_AI_Driveway\dataset\valid_tuned\labels... 91 images, 0 backgrounds, 0 corrupt: 100%[██████████] 91/91 [00:00<00:00, 1679.73it/s]
val: New cache created: C:\Users\fmula\Documents\Yoobee\Subjects\Term_2\Intelligent_TS\Assessment_2\ITS2_AI_Driveway\dataset\valid_tuned\labels.cache
Plotting labels to runs\detect\train2\labels.jpg...
optimizer: 'optimizer=auto' found, ignoring 'lr=0.001' and 'momentum=0.937' and determining best 'optimizer', 'lr=0' and 'momentum' automatically...
optimizer: AdamW(lr=0.000118, momentum=0.9) with parameter groups 69 weight(decay=0.0), 76 weight(decay=0.0005), 75 bias(decay=0.0)
Image sizes 640 train, 640 val
Using 0 dataloader workers
Logging results to runs\detect\train2
Starting training for 50 epochs...
```

Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
1/50	0G	1.432	4.738	1.547	43	640: 100%[██████████] 24/24 [01:22<00:00, 3.44s/it]
	Class	Images	Instances	Box(P	R	mAP50
	all	91	212	0	0	mAP50-95: 100%[██████████] 3/3 [00:13<00:00, 4.64s/it]

Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
34/50	0G	1.04	1.174	1.309	61	640: 100% [██████████] 24/24 [01:49<00:00, 4.55s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:12<00:00, 4.08s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.847	0.926	0.958 0.684
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
35/50	0G	1.044	1.18	1.322	53	640: 100% [██████████] 24/24 [02:10<00:00, 5.45s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:11<00:00, 3.89s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.915	0.819	0.95 0.683
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
36/50	0G	1.056	1.183	1.333	46	640: 100% [██████████] 24/24 [01:50<00:00, 4.61s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:11<00:00, 3.79s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.89	0.911	0.959 0.691
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
37/50	0G	1.043	1.179	1.309	77	640: 100% [██████████] 24/24 [02:02<00:00, 5.09s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:15<00:00, 5.10s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.877	0.924	0.958 0.687
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
38/50	0G	1.058	1.165	1.319	52	640: 100% [██████████] 24/24 [02:35<00:00, 6.50s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:15<00:00, 5.06s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.865	0.901	0.954 0.686
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
39/50	0G	1.039	1.141	1.297	61	640: 100% [██████████] 24/24 [02:31<00:00, 6.31s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:14<00:00, 4.96s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.891	0.909	0.959 0.692
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
40/50	0G	1.039	1.164	1.299	54	640: 100% [██████████] 24/24 [02:32<00:00, 6.35s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:18<00:00, 6.15s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.901	0.905	0.961 0.688

Closing dataloader mosaic

Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
41/50	0G	0.9469	1.253	1.276	23	640: 100% [██████████] 24/24 [02:26<00:00, 6.11s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:14<00:00, 4.77s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.824	0.92	0.958 0.685
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
42/50	0G	0.9007	1.115	1.213	23	640: 100% [██████████] 24/24 [02:29<00:00, 6.21s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:15<00:00, 5.13s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.788	0.945	0.951 0.675
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
43/50	0G	0.8779	1.098	1.289	25	640: 100% [██████████] 24/24 [02:25<00:00, 6.04s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:16<00:00, 5.50s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.927	0.85	0.961 0.673
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
44/50	0G	0.8701	1.086	1.183	24	640: 100% [██████████] 24/24 [02:25<00:00, 6.08s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:12<00:00, 4.24s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.905	0.883	0.961 0.671
Epoch	GPU_mem	box_loss	cls_loss	dfl_loss	Instances	Size
45/50	0G	0.8959	1.076	1.214	28	640: 100% [██████████] 24/24 [02:15<00:00, 5.66s/it] mAP50 mAP50-95: 100% [██████████] 3/3 [00:18<00:00, 6.12s/it]
	Class	Images	Instances	Box(P R		
	all	91	212	0.909	0.891	0.958 0.67

EarlyStopping: Training stopped early as no improvement observed in last 6 epochs. Best results observed at epoch 39, best model saved as best.pt.

To update EarlyStopping(patience=6) pass a new patience value, i.e. 'patience=300' or use 'patience=0' to disable EarlyStopping.

45 epochs completed in 1.577 hours.

Optimizer stripped from runs\detect\train2\weights\last.pt, 5.6MB

Optimizer stripped from runs\detect\train2\weights\best.pt, 5.6MB

Validating runs\detect\train2\weights\best.pt...

Ultralytics 8.3.160 Python-3.13.5 torch-2.7.1+cpu CPU (13th Gen Intel Core(TM) i5-1334U)

YOLOv5n summary (fused): 84 layers, 2,658,071 parameters, 0 gradients, 7.8 GFLOPs

Class	Images	Instances	Box(P	R	mAP50	mAP50-95: 100%
all	91	212	0.891	0.909	0.958	0.692
person	60	61	0.852	0.951	0.955	0.722
kids	91	151	0.93	0.868	0.963	0.662

Speed: 3.5ms preprocess, 147.8ms inference, 0.0ms loss, 2.4ms postprocess per image

Results saved to runs\detect\train2

```
In [9]: # 4. Training complete
print("Training complete. Results:")
print(results)
```

```
Training complete. Results:  
ultralytics.utils.metrics.DetMetrics object with attributes
```


03_load_test_finetune.ipynb

Objective

This notebook aims to load the fine-tuned YOLOv5n model (`best.pt`), which was retrained locally to include a new class ('kids') in addition to the original 80 classes. The fine-tuning process attempted to preserve the previous knowledge by freezing the initial layers of the model.

The original YOLOv5n pre-trained model (80 classes) was downloaded from Ultralytics. For this fine-tuning stage, a new class ('kids') was added, expanding the model's detection capabilities while retaining the original learned features.

Steps:

1. Import the necessary libraries
2. Load the fine-tuned YOLOv5n model
3. Test the model on sample images and videos
4. Display the results

```
In [1]: # 1. Import required libraries
import torch
from pathlib import Path
import cv2
from matplotlib import pyplot as plt
from ultralytics import YOLO

In [2]: # 2. Load the re-trained YOLOv5n new model
model_path = Path("./best.pt")
assert model_path.is_file(), f"Model not found at {model_path}"

# Load the model using ultralytics API
model_ft = YOLO(str(model_path))

In [12]: # 3. Test the model on a sample image
# For this example, place an image named "test_image.jpg" in the same folder
img_path = Path("./img1.jpg")
assert img_path.is_file(), f"Test image not found at {img_path}"

# Run inference
results = model_ft(str(img_path))

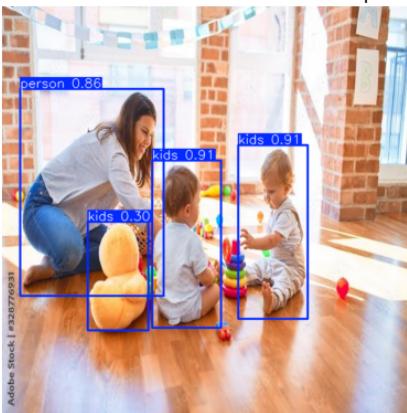
image 1/1 c:\Users\fmula\Documents\Yoobee\Subjects\Term_2\Intelligent_TS\Assessment_2\ITS2_AI_Driveway\img1.jpg: 640x640 1 person, 3 kids, 100.1ms
Speed: 3.1ms preprocess, 100.1ms inference, 1.5ms postprocess per image at shape (1, 3, 640, 640)
```

```
In [19]: import matplotlib.pyplot as plt

img_bgr = results[0].plot()
img_rgb = cv2.cvtColor(img_bgr, cv2.COLOR_BGR2RGB)

plt.imshow(img_rgb)
plt.axis('off')
plt.title("YOLOv5n Detection fine-tuned- best.pt")
plt.show()
```

YOLOv5n Detection fine-tuned- best.pt



```
In [4]: import matplotlib.pyplot as plt
import cv2

# 4. Display the results
results[0].show() # This will open a window showing the image with detections
```

```
In [5]: # You can also access detailed results (e.g., labels, bounding boxes)
df = results[0].to_df()
print(df) # pandas DataFrame with bounding boxes and confidence scores
```

	name	class	confidence
0	kids	80	0.90854
1	kids	80	0.90816
2	person	0	0.86425
3	kids	80	0.30237


```
box
0 {'x1': 370.56799, 'y1': 217.95358, 'x2': 479.2...
1 {'x1': 235.07672, 'y1': 241.98689, 'x2': 343.6...
2 {'x1': 28.29343, 'y1': 129.6394, 'x2': 253.564...
3 {'x1': 134.7287, 'y1': 337.99515, 'x2': 229.13...
```

04_demo_alerts.ipynb

Objective

This notebook is designed to load and evaluate the pre-trained YOLOv5n model (yolov5n.pt) and fine-tuned model (best.pt), which has already been downloaded locally. The goal is to test its inference capabilities on a selection of sample images and videos.

The YOLOv5n model, pre-trained on the COCO dataset (comprising 80 classes), was obtained from Ultralytics. For the current stage of the project, only a subset of relevant object classes has been retained for detection purposes: ['person', 'cat', 'dog', 'horse', 'sheep', 'cow'].

The best model, was fine-tuned on a pre-trained on the COCO dataset adding a new class "kids" (comprising 81 classes). For the current stage of the project, only a subset of relevant object classes has been retained for detection purposes: ['person', 'cat', 'dog', 'horse', 'sheep', 'cow', 'kids'].

Steps:

1. Import required libraries and dependencies
2. Load the pre-trained YOLOv5n e best fine-tuned models
3. Check and filter relevant object classes for both models
4. Run inference on test images and video samples
5. Display detection results and verify alerts

```
In [2]: from ultralytics import YOLO
import cv2
import tkinter as tk
from tkinter import messagebox
import threading
import pyaudio
import wave
from pathlib import Path
```

```
In [3]: # Checkin all classes for each model
# Load model YOLOv5n pre-trained with 80 labels
model = YOLO("yolov5nu.pt")

# Show Labels
for i, label in model.names.items():
    print(f"{i}: {label}")
```

```
0: person
1: bicycle
2: car
3: motorcycle
4: airplane
5: bus
6: train
7: truck
8: boat
9: traffic light
10: fire hydrant
11: stop sign
12: parking meter
13: bench
14: bird
15: cat
16: dog
17: horse
18: sheep
19: cow
20: elephant
21: bear
22: zebra
23: giraffe
24: backpack
25: umbrella
26: handbag
27: tie
28: suitcase
29: frisbee
30: skis
31: snowboard
32: sports ball
33: kite
34: baseball bat
35: baseball glove
36: skateboard
37: surfboard
38: tennis racket
39: bottle
40: wine glass
41: cup
42: fork
43: knife
44: spoon
45: bowl
46: banana
47: apple
48: sandwich
49: orange
50: broccoli
51: carrot
52: hot dog
53: pizza
54: donut
55: cake
56: chair
57: couch
58: potted plant
59: bed
60: dining table
61: toilet
62: tv
63: laptop
64: mouse
65: remote
66: keyboard
67: cell phone
68: microwave
69: oven
70: toaster
71: sink
72: refrigerator
73: book
74: clock
75: vase
76: scissors
77: teddy bear
78: hair drier
79: toothbrush
```

```
In [ ]: # Checkin all classes for each model
# Load model YOLOv5n pre-trained + fine-tune with 81 Labels
model = YOLO("best.pt")

# Show Labels
for i, label in model.names.items():
    print(f"{i}: {label}")
```

```
0: person
1: bicycle
2: car
3: motorcycle
4: airplane
5: bus
6: train
7: truck
8: boat
9: traffic light
10: fire hydrant
11: stop sign
12: parking meter
13: bench
14: bird
15: cat
16: dog
17: horse
18: sheep
19: cow
20: elephant
21: bear
22: zebra
23: giraffe
24: backpack
25: umbrella
26: handbag
27: tie
28: suitcase
29: frisbee
30: skis
31: snowboard
32: sports ball
33: kite
34: baseball bat
35: baseball glove
36: skateboard
37: surfboard
38: tennis racket
39: bottle
40: wine glass
41: cup
42: fork
43: knife
44: spoon
45: bowl
46: banana
47: apple
48: sandwich
49: orange
50: broccoli
51: carrot
52: hot dog
53: pizza
54: donut
55: cake
56: chair
57: couch
58: potted plant
59: bed
60: dining table
61: toilet
62: tv
63: laptop
64: mouse
65: remote
66: keyboard
67: cell phone
68: microwave
69: oven
70: toaster
71: sink
72: refrigerator
73: book
74: clock
75: vase
76: scissors
77: teddy bear
78: hair drier
79: toothbrush
80: kids
```

```
In [ ]: # Load model pre-trained to detect objects
import torch

model = YOLO("yolov5n.pt")
#model = YOLO("best.pt")

audio = 'beep.wav'

video = "video1.mp4" # choose the video name here to test more videos

# Load video
cap = cv2.VideoCapture(video)
frame_width = int(cap.get(cv2.CAP_PROP_FRAME_WIDTH))
frame_height = int(cap.get(cv2.CAP_PROP_FRAME_HEIGHT))

# Parameters
threshold_frames = 3
confidence_threshold = 0.45
consecutive_frames = 0
keep_beeping = False
alert_active = False # NEW: flag to avoid multiple alerts

# Target labels to detect
target_labels = ['person', 'kids', 'cat', 'dog', 'horse', 'sheep', 'cow']

# Function: play alarm sound in loop
def beep_loop():
    global keep_beeping
    chunk = 1024

    wf = wave.open(audio, 'rb')
    p = pyaudio.PyAudio()

    stream = p.open(format=p.get_format_from_width(wf.getsampwidth()),
                    channels=wf.getnchannels(),
                    rate=wf.getframerate(),
                    output=True)
```

```

while keep_beeping:
    wf.rewind()
    data = wf.readframes(chunk)
    while data and keep_beeping:
        stream.write(data)
        data = wf.readframes(chunk)

    stream.stop_stream()
    stream.close()
    p.terminate()

# Function: show popup alert with sound
def show_alert(label, frames_detected):
    global keep_beeping, alert_active
    keep_beeping = True

    # Start beep in thread
    threading.Thread(target=beep_loop, daemon=True).start()

    # Show popup
    root = tk.Tk()
    root.withdraw()
    print(f"ALERT: {label.upper()} detected for {frames_detected} consecutive frames!")
    messagebox.showwarning(
        "DETECTION ALERT",
        f"ALERT: {label.upper()} detected for {frames_detected} consecutive frames!"
    )

    # Stop beep when popup closed
    keep_beeping = False
    root.destroy()

    # Reset alert flag
    alert_active = False

def center_window(window_name, frame_width, frame_height):
    # Get size of the screen
    root = tk.Tk()
    screen_width = root.winfo_screenwidth()
    screen_height = root.winfo_screenheight()
    root.destroy()

    # Calculating the position to center the window
    x = int((screen_width - frame_width) / 2)
    y = int((screen_height - frame_height) / 2)

    # Moving the window to the center
    cv2.moveWindow(window_name, x, y)

window_name = "Detection"

# Create window with normal size and center it
cv2.namedWindow(window_name, cv2.WINDOW_NORMAL)
center_window(window_name, frame_width, frame_height)

# Main Loop
while True:
    ret, frame = cap.read()
    if not ret:
        break

    # Run detection
    results = model(frame, verbose=False)
    annotated = results[0].plot()

    # Check for target labels
    detected = False
    detected_label = ""

    for box in results[0].boxes:
        conf = box.conf.item()
        label = results[0].names[int(box.cls)]

        if label.lower() in target_labels and conf > confidence_threshold:
            detected = True
            detected_label = label
            break

    # Update counter
    consecutive_frames = consecutive_frames + 1 if detected else 0

    # Trigger alert (if not already active)
    if consecutive_frames >= threshold_frames and not alert_active:
        alert_active = True
        threading.Thread(target=show_alert, args=(detected_label, consecutive_frames), daemon=True).start()
        consecutive_frames = 0

    # Show video with detections
    cv2.imshow("Detection", annotated)
    if cv2.waitKey(1) & 0xFF == ord("q"):
        break

# Cleanup
cap.release()
cv2.destroyAllWindows()

# press "q" to finish the video

```

ALERT: KIDS detected for 4 consecutive frames!
 ALERT: KIDS detected for 5 consecutive frames!
 ALERT: KIDS detected for 5 consecutive frames!