

Machine Learning Final Project Proposal - 2023

Project Proposal Outline: Object Detection and Tracking in SUSTech Campus

0. Team Members

- 罗皓予 12112517
- 李昱纬 12112513
- 莫羡瑜 12112712
- 李田 12112628

1. Background

Significance of the Project:

Autonomous driving is at the forefront of technological advancements, aiming to revolutionize transportation by minimizing human intervention. Object detection and tracking serve as the eyes of intelligent vehicles, enabling them to perceive and respond to their surroundings. In the context of South University of Science and Technology (SUSTech) campus, the implementation of robust object detection is essential for navigating the unique challenges posed by campus traffic scenarios.

The SUSTech campus, characterized by diverse pedestrian movement, various vehicle types, and intricate road structures, demands a tailored approach to object detection. Traditional models may struggle to adapt to the nuances of campus environments, making it imperative to explore and enhance existing algorithms for more accurate and efficient detection.

Technological Framework:

Our project is anchored in the utilization of YOLOv8, a cutting-edge algorithm renowned for its real-time object detection capabilities. YOLOv8 stands out due to its ability to process images in a single forward pass, making it highly suitable for applications where low latency is crucial, such as autonomous vehicles. By building upon the YOLOv8 architecture, we aim to customize and optimize the model to excel in the specific challenges presented by SUSTech's dynamic campus setting.

The choice of YOLOv8 is motivated by its versatility and proven success in various computer vision tasks. Leveraging this technology, we seek to contribute to the advancement of autonomous driving within the unique context of a university campus, addressing the need for precise object detection and tracking in dynamic and densely populated environments.

In summary, our project addresses the gap in object detection for autonomous vehicles on the SUSTech campus, leveraging the power of YOLOv8 to provide a solution tailored to the distinctive challenges presented by campus traffic scenarios. This endeavor not only enhances the safety and efficiency of autonomous vehicles within the campus but also contributes valuable insights to the broader field of computer vision and autonomous driving technology.

2. Motivation

Focus Areas for Improvement:

1. Enhancing Detection Accuracy:

- To ensure the safety of autonomous vehicles and campus inhabitants, elevating detection accuracy in the diverse traffic of SUSTech campus is critical.
- Enhanced accuracy is vital for reliable navigation amidst pedestrians and cyclists.

2. Tackling Campus Environmental Challenges:

- The SUSTech campus poses unique challenges such as varied lighting, diverse pedestrian and vehicle movements, requiring specialized detection strategies.
- Adapting detection models to these specific campus conditions is crucial for effective performance.

3. Improving Small-Size Object Detection:

- Detecting small-sized objects, often overlooked by standard models, is essential in the densely populated and intricate campus setting.
- Refining detection capabilities for smaller objects will significantly contribute to overall safety and efficiency.

4. Assessing YOLOv8's Versatility:

- YOLOv8's effectiveness in diverse settings needs exploration, particularly in dynamic environments like university campuses.
- Evaluating YOLOv8's adaptability to the specific needs of campus environments is essential for its broad application.

Core Project Motivations:

This project is driven by the aspiration to advance autonomous driving technology at SUSTech, with a focus on:

- **Increasing Safety:** By improving accuracy, especially in detecting smaller objects, the project aims to ensure safer interactions within the campus.
- **Addressing Specific Campus Challenges:** Customizing detection models to meet the unique demands of campus environments enhances their effectiveness.
- **Expanding YOLOv8's Applications:** Adapting YOLOv8 for a university setting offers insights into its versatility and potential for wider use.

In summary, the project seeks to refine autonomous driving in complex campus environments, contributing to the broader evolution of autonomous vehicle technologies.

3. Related Work

Advances in Object Detection Techniques:

1. Feature Extraction Techniques in Object Detection:

- **Convolutional Neural Networks (CNNs):** Essential for extracting hierarchical features from images, CNNs form the backbone of many object detection models.
- **Pooling Methods:** Max pooling and average pooling are crucial for reducing the spatial size of feature maps, thereby reducing computational load and overfitting.

2. Multi-scale Feature Fusion Approaches:

- **Feature Pyramid Network (FPN):** FPN enhances object detection by integrating low-level detail with high-level semantic information across different scales.
- **Single Shot MultiBox Detector (SSD):** SSD is known for its ability to handle multiple object sizes efficiently, thanks to its multi-scale feature maps.

3. Detection Techniques in Object Detection Models:

- **Non-Maximum Suppression (NMS):** NMS plays a pivotal role in eliminating redundant bounding boxes, ensuring only the most probable one is selected for each object.
- **Sliding Window Approach:** This technique involves moving a window across the image and evaluating it at each position, a fundamental concept in earlier detection methods.

4. YOLO Network Architecture and Evolution:

- **YOLO Architecture:** YOLO's revolutionary design processes the entire image in a single forward pass, enabling fast and efficient object detection.
- **Evolution to YOLOv8:** YOLOv8 builds on this foundation, further improving accuracy and speed, making it particularly relevant for real-time applications like autonomous driving.

Implications for Campus-Specific Applications:

1. Customization for Dynamic Campus Environments:

- **Adapting to Diverse Campus Settings:** The dynamic and varied nature of university campuses, like SUSTech, presents unique challenges that standard models may not address effectively.
- **Enhancing Small Object Detection:** Recognizing the importance of detecting smaller objects in dense campus environments, where they are often overlooked, is crucial.

2. Leveraging YOLOv8 for Campus Applications:

- **YOLOv8's Potential in Campus Scenarios:** While YOLOv8 has shown success in various settings, its effectiveness in the specific context of a campus, with its unique traffic and pedestrian dynamics, needs exploration.
- **Incorporating Campus-Specific Features:** There is scope to improve object detection on campuses by integrating features unique to these environments, such as varied pedestrian activities and complex traffic patterns.

4. Novelty of This Work

Enhancements in Object Detection Techniques:

1. Customizing YOLOv8 for Campus Environments:

- *Targeted Modifications:* Adaptations are made to YOLOv8 to suit SUSTech's unique campus dynamics, focusing on diverse pedestrian behavior and variable traffic conditions.
- *Campus-Centric Optimization:* The model is fine-tuned to improve detection in pedestrian-dense areas and under varied lighting, enhancing accuracy and relevance.

2. Refining Detection Accuracy and Efficiency:

- *Precision Parameter Tuning:* Parameters are meticulously adjusted based on campus-specific data, aiming to increase both detection precision and recall.
- *Reducing False Positives:* Special attention is given to challenging scenarios like pedestrian crossings, aiming to minimize false detections for greater system reliability.

3. Extensive Real-World Testing and Validation:

- *Campus-Based Evaluation:* Although YOLOv8 is effective in various settings, its performance in university campuses is tested and validated, specifically in the SUSTech context.
- *Insights for Broader Application:* Findings from this project are expected to aid in developing tailored object detection solutions for other university campuses.

4. Integrating Transformer Self-Attention with Linear Transformers

Advanced Feature Recognition:

- **Self-Attention Mechanism:** We integrate a transformer self-attention mechanism into the YOLOv8 architecture. This innovation allows the model to focus on relevant parts of the image by weighing the importance of different features. It's particularly effective in complex scenes where traditional convolutional layers might miss subtle but crucial details.
- **Contextual Understanding:** The self-attention mechanism aids in better contextual understanding of the scene, enhancing the model's ability to differentiate between closely situated objects and recognize patterns in dynamic environments, such as a bustling university campus.

Efficiency and Speed Optimization:

- **Linear Transformers:** To address the typically high computational demands of transformer models, we utilize linear transformers. These are designed to reduce the computational complexity from quadratic to linear, enabling faster processing without compromising the depth of feature analysis.
- **Real-Time Processing Capability:** The incorporation of linear transformers ensures that the enhanced feature recognition capabilities of the self-attention mechanism do not adversely impact the real-time processing requirements essential for autonomous driving applications.

Enhanced Object Detection:

- **Improved Accuracy in Dense Environments:** By employing the self-attention mechanism, our model becomes more adept at detecting small or partially obscured objects in densely populated areas, a common challenge in campus settings.
- **Adaptability to Varied Scenarios:** The combination of self-attention and linear transformers allows the model to adapt more effectively to varied scenarios encountered on a university campus, such as rapidly changing pedestrian flows, diverse vehicle types, and fluctuating lighting conditions.

5. Methodology

Approach Components:

1. YOLOv8 as the Base Model:

- Adopting YOLOv8 for its robust architecture, serving as the core framework.
- Employing its efficient feature extraction and prediction in a single forward pass, essential for real-time applications.

2. Integration of Campus-Specific Features:

- Tailoring the model to address unique aspects of the SUSTech campus, such as diverse pedestrian movements and complex road layouts.
- Adapting to various environmental factors like lighting conditions and traffic diversity on campus.

3. Detailed Parameter Adjustment:

- Fine-tuning model parameters to enhance precision and adaptability in campus settings.
- Concentrating on refining aspects critical for object detection, like threshold settings, anchor box dimensions, and improved suppression methods.

4. Incorporating Transformer Self-Attention Mechanisms:

- Implementing transformer-based self-attention mechanisms to improve the model's capability in recognizing and differentiating objects in complex scenes.
- Aiming to enhance the model's contextual understanding, particularly beneficial in crowded or dynamically changing campus environments.

6. Resources and Experiment Platform

Data Collection:

1. Image Source:

- Utilizing a diverse set of traffic images captured within the SUSTech campus.
- Ensuring images cover various scenarios, including crowded pedestrian areas, vehicle interactions, and different lighting conditions.

2. Annotation Process:

- Manually annotating images to create a labeled dataset for training and evaluation.
- Emphasizing precise labeling for objects such as pedestrians, vehicles, and cyclists.

3. Dataset Augmentation:

- Introducing variations in scale, rotation, and lighting to enhance model generalization.
- Augmenting the dataset to simulate diverse campus scenarios not captured in the original images.

Experiment Platform:

1. Hardware Infrastructure:

- **GPU Resources:** Employing high-performance GPUs (e.g., NVIDIA Tesla V100) to expedite training and inference processes.
- **Storage:** Utilizing sufficient storage capacity for storing large datasets and model checkpoints.

2. Software Dependencies:

- **Deep Learning Framework:** Leveraging PyTorch as the primary deep learning framework for model development and training.
- **YOLOv8 Codebase:** Utilizing the YOLOv8 open-source implementation as the baseline for modifications.
- **Python Libraries:** Employing essential libraries for image processing, data augmentation, and performance evaluation.

3. Collaboration Tools:

- Using collaborative platforms such as GitHub for version control and team collaboration.
- Implementing communication channels for real-time collaboration and issue tracking.

- **Experimentation Workflow:**

1. Data Collection for Training and Testing:

- **Gathering Diverse Data:** Collecting a wide range of images and videos from the SUSTech campus environment, ensuring a variety of scenarios including different times of day, weather conditions, and pedestrian and vehicle densities.
- **Annotation and Preprocessing:** Annotating this data meticulously to label various objects like pedestrians, cyclists, and vehicles. Also preprocessing the data for optimal model input, such as resizing, normalization, and augmentation techniques.

2. Training Phase:

- **Model Training:** Iteratively training the enhanced YOLOv8 model with the annotated dataset, focusing on the integration of transformer self-attention mechanisms.
- **Parameter Optimization:** Fine-tuning various parameters like learning rate, batch size, and architecture-specific settings based on performance feedback and intermediate evaluation results.

3. Testing and Validation:

- **Simulated Scenario Testing:** Conducting tests using simulated campus scenarios to assess the model's preliminary performance.
- **Real-World Validation:** Implementing extensive testing in actual campus environments to validate the model's effectiveness in real-world conditions and varied situations.

4. Feedback and Iterative Refinement:

- **Performance Review:** Analyzing testing results and gathering feedback from potential users and stakeholders to evaluate the model's practical applicability and user experience.
- **Iterative Improvements:** Continuously refining the model based on feedback, test results, and evolving requirements, aiming for ongoing optimization and enhancement.

5. Deployment Readiness Assessment:

- **Deployment Trials:** Conducting trial runs in controlled campus environments to assess the readiness of the model for deployment in autonomous vehicles.
- **Final Evaluations:** Finalizing the model post-assessment, ensuring it meets the required safety, accuracy, and efficiency standards for autonomous driving in a campus setting.

7. Goals and Objectives

Core Goals:

Performance Improvement:

- **Objective:** Surpass YOLOv8's performance in our custom campus dataset.
- **Metric:** Achieve better accuracy rates than baseline YOLOv8 in identifying campus-specific scenarios.

Evaluation Plan:

Comparative Testing:

- **Benchmarking:** Compare our modified YOLOv8 model against the original YOLOv8 on the campus dataset.
- **Focus Areas:** Assess improvements in accurately detecting pedestrians, vehicles, and other campus-specific elements.

Success Criteria:

Quantitative Measures:

- **Accuracy Target:** Aim for a minimum 3% improvement in overall accuracy compared to the baseline YOLOv8 model.
- **Specific Scenarios:** Excel in campus-specific tests, such as pedestrian crossings and diverse traffic conditions.

Key Milestones:

Model Development and Testing:

- **Initial Model Completion:** Finalize the first iteration of the modified YOLOv8 model.
- **Preliminary Testing:** Conduct initial tests and compare results with baseline YOLOv8.

Acceptance Standards:

Model Validation:

- **Criteria:** Accept the model if it consistently outperforms the baseline YOLOv8 in campus-specific tests.
- **User Feedback:** Incorporate user insights to ensure the model's practicality and effectiveness in real scenarios.

Overall Significance:

Impact on Autonomous Driving:

- **Autonomous Vehicle Enhancement:** Highlight the project's role in advancing object detection for autonomous vehicles in campus environments.
- **Research Contribution:** Emphasize its value to academic research, setting new standards for object detection in specialized settings.

Risk Management:

Proactive Strategies:

- **Progress Tracking:** Regularly evaluate progress against goals and adapt strategies as needed.
- **Feedback Incorporation:** Continuously integrate user and stakeholder feedback to refine the model.

This structured approach, focused on surpassing YOLOv8's performance with tailored enhancements for campus environments, is designed to ensure the successful development and validation of an optimized object detection model for autonomous vehicles.

8. Project Timeline and Responsibilities

Team Roles and Duties:

1. Machine Learning Engineer (MLE):

- Responsibilities:
 - Leading the development and optimization of the YOLOv8 model.
 - Conducting training, testing, and iterative improvements based on results.

2. Data Scientist (DS):

- Responsibilities:
 - Collecting and annotating data from the SUSTech campus.
 - Ensuring the dataset is comprehensive and ready for model training.

Detailed Schedule:

Week 1: Data Collection and Preparation

- DS:
 - Completes comprehensive data collection from the SUSTech campus.
 - Begins the process of data annotation and preparation for training.
- MLE:
 - Assists with data collection strategies and preparation guidelines.

Week 2: Model Implementation and Setup

- MLE:
 - Finalizes the YOLOv8 network modifications and sets up the training environment.
- DS:
 - Continues with detailed annotation and dataset finalization.

Week 3: Model Training and Initial Testing

- MLE:
 - Commences training of the model using the prepared dataset.
 - Conducts initial tests to assess model performance and identify areas for improvement.
- DS:
 - Supports the MLE with additional data requirements and adjustments based on initial testing results.

Week 4: Final Testing and Optimization

- MLE:
 - Carries out extensive testing and fine-tuning of the model.
 - Analyzes test results for final improvements.
- DS:
 - Assists in the final testing phase by providing additional data insights and validation support.

Key Milestones:

- **End of Week 1:** Data collection and initial annotation complete.
- **End of Week 2:** Network implementation and training environment ready.
- **End of Week 3:** Initial model training and testing conducted.
- **End of Week 4:** Final model optimization and testing completed.

Collaboration Strategy:

- **Regular Synchronization:** MLE and DS maintain frequent communication to synchronize efforts and share insights.
- **Flexible Task Sharing:** Both team members are prepared to share tasks and responsibilities as needed to meet the tight timeline.
- **Progress Monitoring:** Continuous monitoring of progress and readiness to adjust strategies based on the results.

This streamlined approach focuses on efficient and effective collaboration between the MLE and DS, ensuring the project's success within the one-month timeframe. Regular communication and flexibility in task sharing are key to navigating the challenges and meeting the project's objectives.