Project 2350 | October 2024

SJSU SAN JOSÉ STATE UNIVERSITY

MTI MINETA TRANSPORTATION INSTITUTE

Artificial Intelligence for Pedestrian and Bicyclist Safety: Using AI to Detect Near-Miss Collisions

Mohammad Pourhomayoun, PhD

CSUTC California State University Transportation Consortium

CAL STATE LA
CALIFORNIA STATE UNIVERSITY, LOS ANGELES

MINETA TRANSPORTATION INSTITUTE

Founded in 1991, the Mineta Transportation Institute (MTI), an organized research and training unit in partnership with the Lucas College and Graduate School of Business at San José State University (SJSU), increases mobility for all by improving the safety, efficiency, accessibility, and convenience of our nation's transportation system. Through research, education, workforce development, and technology transfer, we help create a connected world. MTI leads the Mineta Consortium for Transportation Mobility (MCTM) and the Mineta Consortium for Equitable, Efficient, and Sustainable Transportation (MCEEST) funded by the U.S. Department of Transportation, the California State University Transportation Consortium (CSUTC) funded by the State of California through Senate Bill 1 and the Climate Change and Extreme Events Training and Research (CCEETR) Program funded by the Federal Railroad Administration. MTI focuses on three primary responsibilities:

Research

MTI conducts multi-disciplinary research focused on surface transportation that contributes to effective decision making. Research areas include: active transportation; planning and policy; security and counterterrorism; sustainable transportation and land use; transit and passenger rail; transportation engineering; transportation finance; transportation technology; and workforce and labor. MTI research publications undergo expert peer review to ensure the quality of the research.

Education and Workforce Development

To ensure the efficient movement of people and products, we must prepare a new cohort of transportation professionals who are ready to lead a more diverse, inclusive, and equitable transportation industry. To help achieve this, MTI sponsors a suite of workforce development and education opportunities. The Institute supports educational programs offered by the Lucas Graduate School of Business: a Master of Science in Transportation Management, plus graduate certificates that include High-Speed and Intercity Rail Management and Transportation Security Management. These flexible programs offer live online classes so that working transportation professionals can pursue an advanced degree regardless of their location.

Information and Technology Transfer

MTI utilizes a diverse array of dissemination methods and media to ensure research results reach those responsible for managing change. These methods include publication, seminars, workshops, websites, social media, webinars, and other technology transfer mechanisms. Additionally, MTI promotes the availability of completed research to professional organizations and works to integrate the research findings into the graduate education program. MTI's extensive collection of transportation-related publications is integrated into San José State University's world-class Martin Luther King, Jr. Library.

Report 24-34

Artificial Intelligence for Pedestrian and Bicyclist Safety: Using AI to Detect Near-Miss Collisions

Mohammad Pourhomayoun, PhD

October 2024

A publication of the
Mineta Transportation Institute
Created by Congress in 1991

College of Business
San José State University
San José, CA 95192-0219

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. 24-34	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle Artificial Intelligence for Pedestrian and Bicyclist Safety: Using AI to Detect Near-Miss Collisions		**5. Report Date** October 2024 **6. Performing Organization Code**
7. Authors Mohammad Pourhomayoun, PhD		**8. Performing Organization Report** CA-MTI-2350
9. Performing Organization Name and Address Mineta Transportation Institute College of Business San José State University San José, CA 95192-0219		**10. Work Unit No.** **11. Contract or Grant No.** ZSB12017-SJAUX
12. Sponsoring Agency Name and Address State of California SB1 2017/2018 Trustees of the California State University Sponsored Programs Administration 401 Golden Shore, 5th Floor Long Beach, CA 90802		**13. Type of Report and Period Covered** **14. Sponsoring Agency Code**
15. Supplemental Notes 10.31979/mti.2024.2350		

16. Abstract

Near-Miss Collisions are events that, with a slight change in position or timing, could have resulted in a collision, which could have caused severe injury or property damage. Understanding near-miss collisions can help identify risks and potentially improve road safety. In this project, we developed an effective end-to-end system based on advanced artificial intelligence (AI) models and computer vision algorithms to detect and report near-miss collisions as an important indicator to identify and measure safety risks, especially in specific circumstances such as a right turn on a red light. The main objective is to improve the safety of pedestrians and bicyclists, by applying automated AI-powered systems to detect accident risks for pedestrians and cyclists. The developed system includes algorithms for detecting and tracking all traffic objects including pedestrians and bicyclists, as well as algorithms for estimating collision risks and detecting near misses. We evaluated the developed system on real videos captured by actual traffic cameras in the city of Los Angeles. Despite the low quality of some of the videos, our results demonstrate high accuracy of the developed models in identifying traffic collision risks and detecting near-misses. The information generated by the developed system allows us to enhance safety measures for pedestrians and bicyclists while simultaneously optimizing traffic flow.

17. Key Words Artificial intelligence, Machine learning, Pedestrian counts, Pedestrian safety, Traffic safety.	**18. Distribution Statement** No restrictions. This document is available to the public through The National Technical Information Service, Springfield, VA 22161.		
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 27	**22. Price**

Form DOT F 1700.7 (8-72)

DOI: 10.31979/mti.2024.2350

Mineta Transportation Institute
College of Business
San José State University
San José, CA 95192-0219

Tel: (408) 924-7560
Fax: (408) 924-7565
Email: mineta-institute@sjsu.edu

transweb.sjsu.edu/research/2350

Acknowledgments

The author would like to thank the Mineta Transportation Institute at San José State University for their support of this project. The author especially appreciates the tremendous support, guidance, and collaboration of Dr. Hilary Nixon, Deputy Executive Director, and Dr. Karen E. Philbrick, Executive Director of the Mineta Transportation Institute. The author also deeply appreciates the work of Dr. Jeanne Holm for her remarkable help and insightful advice throughout the project.

Contents

List of Figures

Executive Summary

"Near-miss" collisions refer to incidents that, with minimal changes in position or timing, could have resulted in a collision, potentially leading to severe injury or property damage. Analyzing these near-miss events is critical for identifying risks and enhancing road safety. In this project, we developed a robust end-to-end system leveraging advanced artificial intelligence (AI) models and computer vision algorithms to detect and report near-miss collisions as a key indicator for assessing and quantifying safety risks. The primary objective is to improve pedestrian and cyclist safety through the deployment of automated AI-driven systems designed to detect potential accident risks. The system incorporates algorithms for detecting and tracking all traffic entities, including pedestrians and bicyclists, along with algorithms for collision risk estimation and near-miss detection. The insights generated by this system enable the enhancement of safety protocols for pedestrians and bicyclists while concurrently optimizing traffic flow management.

1. Introduction

According to the latest reports by the U.S. Department of Transportation, the Insurance Institute for Highway Safety (IIHS), and the National Highway Traffic Safety Administration (NHTSA), the number of traffic fatalities has significantly increased in 2021 and reached a 16-year high [1]-[3]. Traffic fatalities in the state of California have increased by more than 10.7% in 2021 compared to 2020 to 4,258 deaths, which is more than 10 deaths per 100,000 population [2]-[3]. According to the reports, 43% of the victims were pedestrians and cyclists. The city of Los Angeles alone has one of the highest rates of traffic deaths among large U.S. cities. Caltrans and local departments of transportation can optimize their traffic signal system to improve vehicular travel times using the Automated Traffic Surveillance and Control (ATSAC) System for vehicles and Regional Integration of Intelligent Transportation Systems (RIITS). However, understanding the movement of people, bicycles, and their interaction with vehicles is critical to avoid traffic accidents and improve safety. Currently, there is no efficient automated system in the state of California for detecting and predicting collision risks of pedestrians and bicyclists in major urban areas.

According to the Occupational Safety and Health Administration (OSHA), a "near-miss" collision is an incident in which no property was damaged and no personal injury was sustained, but where, given a slight shift in time or position, damage or injury easily could have occurred [4]. Near-miss collisions are dangerous traffic situations that are rarely reported to the authorities. Nevertheless, they are important indicators to identify potential risks and prevent actual collisions with personal injury or property damage in the future. Fortunately, with an effective detection and reporting system, we will be able to take advantage of near-miss collisions to help prevent actual collisions in the future. Such a system will let DOTs, city designers, decision makers, local governments, and other stakeholders understand, detect, and predict collision risks and prevent actual collisions with personal injury or property damage.

With the advancement of technology, automated traffic monitoring has been gaining traction over the past decades. Several methods have been proposed for pedestrian detection in the past couple of years [5]-[9]. These methods have used different techniques including image and video processing algorithms, as well as machine learning/deep learning techniques to detect human targets (pedestrian) in an image or video. Dollar et al. [5] and Beneson et al. [7] put together the most popular and well-annotated pedestrian detection datasets, and evaluated the performance of the most promising pedestrian detectors across several datasets. They have shown that despite significant progress in the past few years, the performance still has much room for improvement especially for low resolution images/videos and for occluded pedestrians in the image.

In a previous project, Pourhomayoun and his research team developed an effective system based on deep learning models and computer vision for pedestrian and bicyclist recognition and counting [8]-[9]. This system achieved the accuracy of more than 96% in pedestrian counting on traffic videos captured by actual traffic cameras in the city of Los Angeles [8]-[9]. This project has been recognized by the USDOT as a remarkable accomplishment and highlighted in the USDOT quarterly newsletter in April 2021 [10]. It was also highlighted in the TRB newsletter under federal research news [11]. This research has also received *The Best Paper Award* in the International Conference on Advances in Signal, Image & Video Processing [8].

Despite significant advancements in computer vision and AI technologies over the past few years, particularly in the detection and counting of pedestrians and bicyclists, the task of collision risk detection and assessment remains a challenging and unresolved problem. The complexity arises from the need to accurately predict and evaluate potential interactions between vehicles and vulnerable road users in dynamic and often unpredictable urban environments. Several factors contribute to the difficulty of developing reliable collision detection systems. First, the diverse nature of urban settings, characterized by varying lighting conditions, occlusions, and high levels of activity, requires AI models to be exceptionally robust and adaptive. Second, human behavior is inherently unpredictable, adding a layer of complexity to accurately forecasting pedestrian and bicyclist movements. Moreover, the integration of real-time data from multiple sources, such as cameras, sensors, and vehicle telematics, demands sophisticated algorithms capable of processing vast amounts of information quickly and accurately.

The main goal of this project was to develop an end-to-end system based on advanced artificial intelligence (AI) models and effective computer vision algorithms to detect and report near-miss collisions as an important indicator to identify and measure safety risks, especially in specific circumstances such as a right turn on a red light. The main objective is to improve the safety of pedestrians and bicyclists by applying automated AI-powered systems to detect accident risks for pedestrians and cyclists.

It is important to notice that in this project we did not install new cameras or add new components or sensors to the traffic operation infrastructure. We used the video streams captured by existing traffic cameras installed and operated by Caltrans or the cities.

2. System Architecture

In this project, we developed an end-to-end system including a series of image/video processing, computer vision algorithms, machine learning, and optimal state estimator and tracking algorithms. The developed system receives traffic videos and monitors, recognizes, and tracks pedestrians, cyclists, cars, trucks, and buses, and then detects the directions and predicts the future locations of each object. Then, the algorithm detects near-miss collisions as situations where a slight shift in time or position of the person or car could have caused an actual collision. The system will also report the near-miss collision detection results.

Figure 1 shows the proposed system architecture. The four main parts of the system include: a) video streaming and preprocessing, b) AI for object detection, c) AI and optimal state estimator for trajectory prediction and tracking, d) risk estimation and near-miss collision detection.

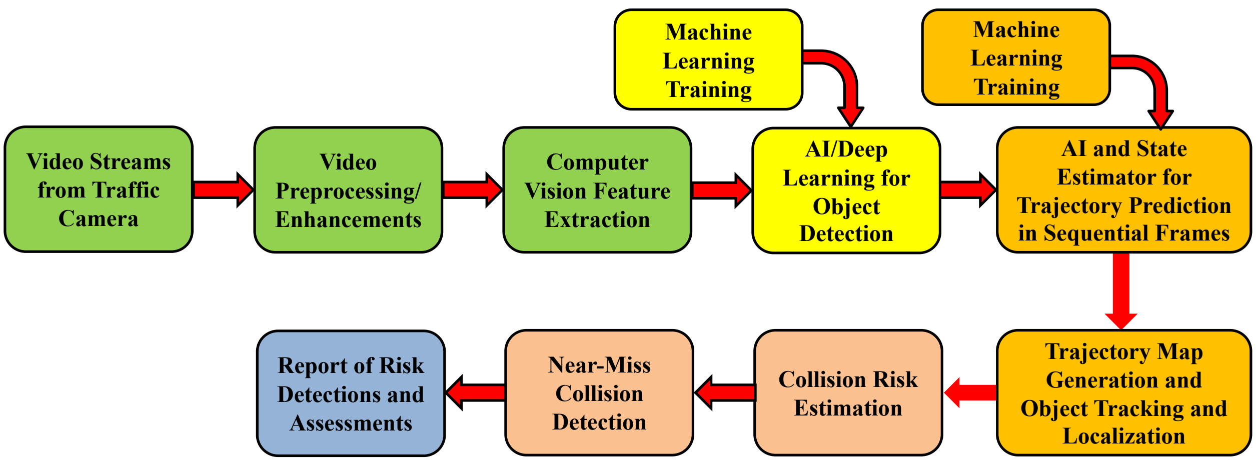


Figure 1. The High-Level System Architecture

3. Data Preprocessing

Implementing traffic computer vision systems in practice, on real-time video streams captured by regular traffic cameras, is substantially different from ideal lab settings. There are challenges and difficulties associated with the type and quality of videos when applying the algorithm to a populated urban area. Some of these challenges include:

- poor quality of video streams due to low resolution, light conditions, disturbed and disoriented lens, or weather conditions;
- dealing with stretched, convex, or squeezed images collected by existing traffic cameras and wide-angle lenses;
- undesired angle, location, and direction of the camera;
- vibration of camera due to the wind or passing of heavy vehicles;
- light distortion at night-time;
- inconsistent lighting during the daytime and shadow effects; and
- moving or stationary objects that may block the view of the target.

Since we intended to work with existing traffic cameras (no need to install new cameras), an important objective of this project was to address the challenges linked to video quality and potentially unfavorable camera locations. The first step in an end-to-end traffic vision system was raw video preprocessing, which includes a series of algorithms for quality enhancement, as well as brightness and contrast adjustment. In the case of wide-angle lenses that may make the image convex, we used regularization algorithms to convert the video back to the normal view. The next step was to identify the area of interest in each frame of the video to improve the performance and accuracy of the object recognition algorithm and reduce the computational load of the system.

4. Machine Learning and Deep Learning for Object Detection

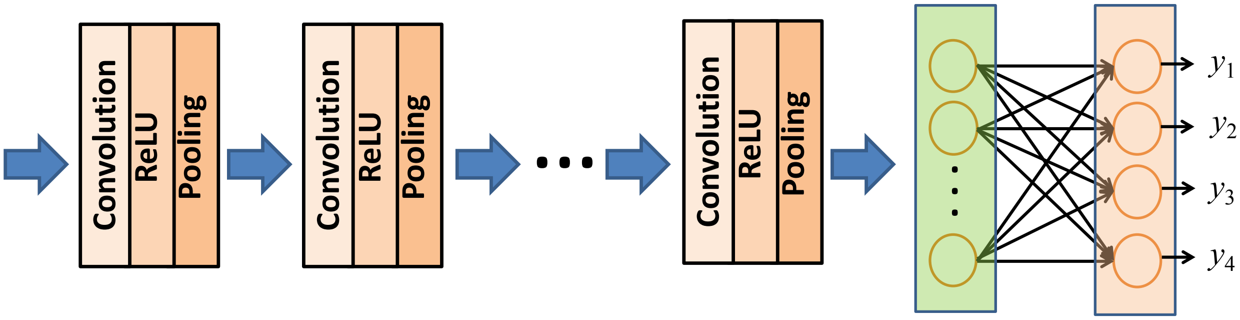
After data preprocessing, the next step was to extract and select the best set of computer vision features that can be used in machine learning algorithms for object detection. Depending on the type of machine learning algorithm, this step included feature extraction, feature selection, and dimensionality reduction. We have tried many different types of features and machine learning algorithms for object recognition [5].

In this project, we have particularly trained and used deep learning models (deep neural network models) including the family of deep Convolutional Neural Networks (ConvNets or CNN), Region-based Convolutional Network (RCNN, Fast RCNN, Faster RCNN), and YOLO (You Only Look Once) algorithms for object detection [8][9][14]-[16].

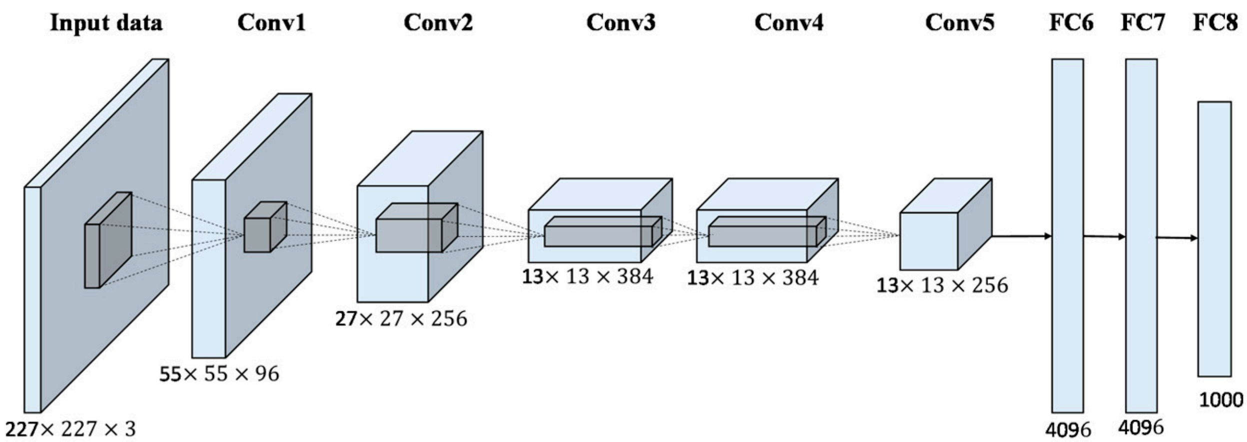
A big advantage of ConvNet methods compared to other classic machine learning algorithms is that there is no need to generate and use hand-made features for ConvNet. The algorithm automatically learns to generate the best set of convolutional features that can best represent the image. However, ConvNet is computationally expensive and sometimes makes it more difficult to run in real-time on high-frame-rate videos. In addition, when the training dataset is not large enough, it is usually hard to train a deep neural network. In previous projects, to reduce the computational complexity, we used to also use traditional machine learning algorithms such as SVM classifiers along with HOG features (Histogram of Oriented Gradient) as an effective method for object recognition with very low computational complexity. However, with new advances in GPU technology, we are now able to train and use very deep neural network models with higher object detection accuracy and still run them in real-time on traffic video streams. Furthermore, *Transfer Learning* methods that take advantage of a pre-trained neural network model on another dataset can be very helpful to ease and expedite the training stage [5].

Figure 2(a) shows the high-level general structure of a ConvNet. Figure 2(b) shows a specific design of ConvNet named AlexNet [15].

Figure 3(a) shows our pedestrian detection results using HOG features and SVM classifier. Figure 3(b) shows our results using YOLO algorithm.

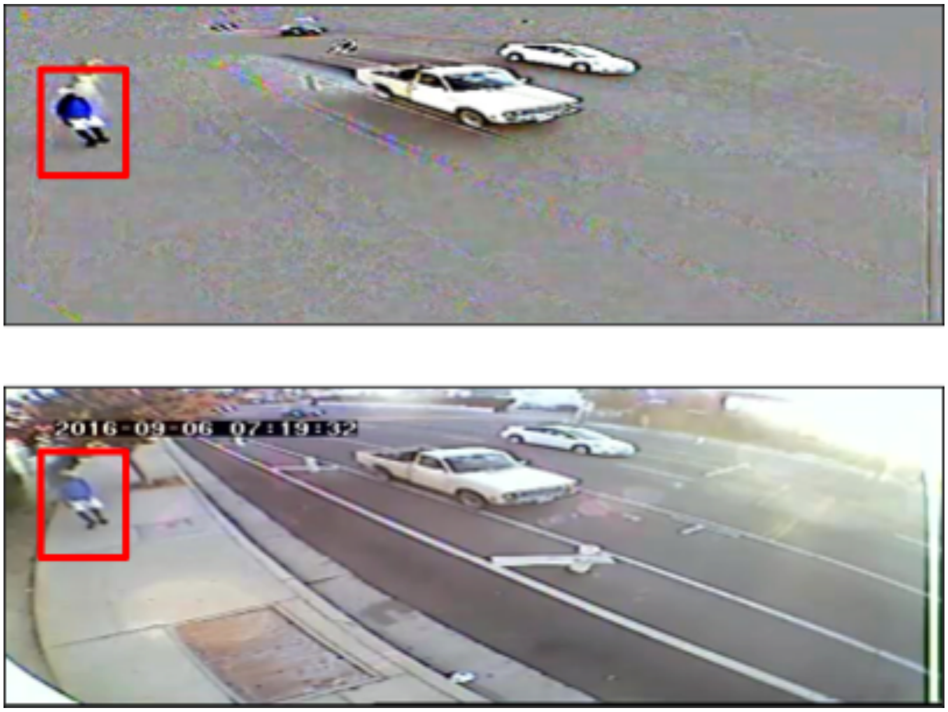


2(a)



2(b)

Figure 2. (a) General Structure of a ConvNet, (b) AlexNet



3(a)

3(b)

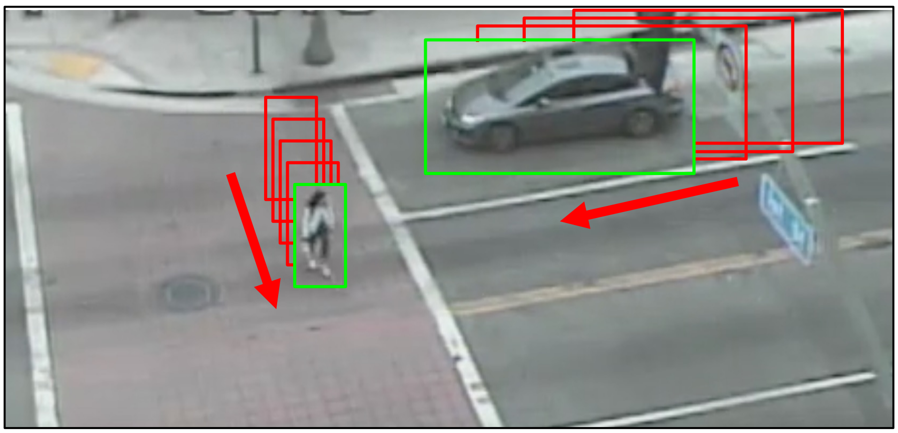
Figure 3. Pedestrian Detection Using Machine Learning Algorithms
(a) Using HOG Features and SVM Classifier, (b) Using YOLO

5. Trajectory Prediction for Traffic Tracking

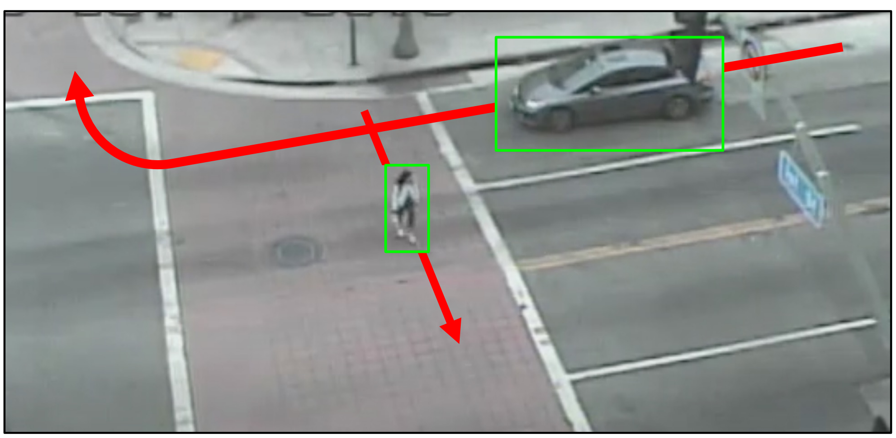
After detecting target objects (e.g., pedestrians, bicyclists, cars, etc.) in several sequential frames, we used Optimal State Estimators to estimate the trajectory of each target object. Since several objects may exist in each frame at a time (e.g., several pedestrians and cars moving in same direction or different directions), it is essential to estimate the trajectory of each object individually. We used Kalman filters and Hungarian models to address data association in sequential frames to estimate the trajectory of the objects over time and predict the next location of each object [17]. This allowed us to track each object individually during the video. In this approach, we used Kalman filters to predict the next location of the moving objects in the next frame based on its previous locations and estimated speed. Then, after receiving the next frame of the video, we compared our prediction results with the actual object detected in next frame, updated the model parameters and estimation (such as the object speed and direction), and continued predicting the next locations in the next frame. Using this approach, we could build a trajectory map including individual trajectories for every object in the video, and then track each object from the first frame it enters until the last frame when it moves out. In this step, the algorithm can also identify all turns including turns on a red light for future risk analysis.

Figure 4 shows our results for (a) object detection, (b) object tracking in sequential frames, and (c) creating a trajectory map and predicting the next location/direction of each moving object.

4(a)



4(b)



4(c)

Figure 4. (a) Object Detection; (b) Object Tracking in Sequential Frames; (c) Creating a Trajectory Map and Predicting the Next Location/Direction of Each Moving Object

6. Near-Miss and Collision Risk Detection

After detecting every traffic object including pedestrians, bicyclists, and all motorized vehicles, the algorithm creates a trajectory map including individual trajectories for every moving object (Figure 4(c)). The trajectory map provides a complete and accurate understanding of object locations along with their directions and speed. It will also let us predict the next location of each moving object.

Machine learning algorithms based on artificial neural networks and maximum likelihood estimation algorithms will be used to classify trajectories into incidents/non-incidents and identify accident risks. The algorithm can detect near-miss collisions as situations where a slight shift in time or position of the person or vehicle could have caused an actual collision. The algorithm detects instances when a moving vehicle is about to collide with a pedestrian or bicyclist. It identifies and rates risks based on distance, velocity, direction, and the types of the motorized vehicles. The models will be trained on traffic videos captured from actual traffic cameras in the city of LA.

Our results demonstrated high accuracy of the proposed models in identifying traffic collision risks and detecting near-misses. Figures 5 and 6 show the sample results of our system on real traffic videos in the city of Los Angeles. Figure 5 shows (a) the object detection algorithm results, (b) risk estimation/assessment based on speed, direction, and distance, and (c) near-miss collision detection (orange circle). Figure 6 shows the case when the algorithm detects a near-miss collision during a “turn on a red light” (orange circle).

5(a)

5(b)

5(c)

Figure 5. (a) Object Detection; (b) Risk Estimation Based on Speed, Direction, and Distance; (c) Near-Miss Collision Detected (Orange Circle)

6(a)

6(b)

Figure 6. (a) Object Detection; (b) Near-Miss Collision Detected During a "Turn on a Red Light" (Orange Circle).

6. Conclusion

Traffic accidents involving pedestrians and cyclists are a significant national concern, with devastating consequences for individuals and communities. These vulnerable road users face a higher risk of severe injuries or fatalities due to their lack of physical protection compared to vehicle occupants. Urban areas in particular present numerous hazards, as dense traffic, inadequate infrastructure, and driver negligence contribute to dangerous conditions. Pedestrians and cyclists often navigate environments where they must share the road with fast-moving vehicles, putting them at risk. In many regions, accidents involving these groups are on the rise, highlighting the urgent need for improved safety measures, such as dedicated bike lanes, enhanced crosswalks, and better enforcement of traffic laws. The integration of advanced technologies like AI and computer vision plays a critical role in addressing traffic safety concerns. AI-powered systems can analyze traffic patterns and facilitate the development of smart infrastructure that prioritizes pedestrian and cyclist safety. Real-time data from AI models can inform urban planners and policymakers, enabling them to design safer city layouts and implement responsive traffic management systems.

The city of Los Angeles has one of the highest rates of traffic deaths among large U.S. cities. Fortunately, the city has adopted the *Vision Zero* initiative as a strategy and commitment to reduce traffic fatalities. Since the most vulnerable components of traffic accidents are pedestrians and bicyclists, it is essential to develop intelligent transportation systems, and human-centered traffic approaches to protect pedestrians and bicyclists.

Near-miss collisions are perilous traffic events that often go unreported to authorities. However, they serve as crucial indicators for identifying potential hazards and preventing actual collisions that could result in personal injury or property damage. By implementing an effective detection and reporting system, we can utilize data from near-miss collisions to avert future accidents. Such a system will enhance our ability to understand, detect, and predict collision risks, enabling us to proactively prevent incidents that could lead to injuries or property damage.

In this project, we have developed an automated system based on advanced AI and computer vision to understand collision risks for pedestrians and bicyclists, particularly to detect and report near-miss collisions as an important indicator to identify and measure actual risks. The real-time information generated by the proposed system allows us to improve safety measures for pedestrians and bicyclists as well as optimizing the flow of traffic and travel time. Despite of many practical challenges (mentioned in chapter 1), the developed system works very well with the existing regular traffic. This system may help increase safety and traffic flow through better traffic management and planning.

Improving the safety of pedestrians and bicyclists has always been a top priority of transportation officials in California. Two of the main goals of the Pedestrian and Bicycle Safety Branch under Caltrans' Traffic Operations are developing programs to improve the safety of transportation infrastructures for pedestrians and bicycles; and encouraging research and technology transfer in the field of pedestrian and bicycle fatality. This project will facilitate these goals by providing a platform to enhance and improve safety measures for pedestrians and bicyclists in California. This project will help transportation authorities advance *Vision Zero* implementation across the state and support the state's effort to reduce traffic fatalities.

Bibliography

1. IIHS, https://www.iihs.org/topics/fatality-statistics/detail/state-by-state.

2. USDOT, https://www.nhtsa.gov/press-releases/early-estimate-2021-traffic-fatalities.

3. USDOT, https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813283.

4. OSHA, https://www.osha.com/blog/near-miss-definition.

5. N. Dalal, B. Triggs: Histograms of oriented gradients for human detection. In: CVPR.

6. P. Dollar, C. Wojek, B. Schiele and P. Perona: "Pedestrian Detection: An Evaluation of the State of the Art", in IEEE Trans. on Pattern Analysis and Machine Intelligence.

7. R. Benenson et al. "Ten Years of Pedestrian Detection, What Have We Learned?" ECCV, Springer.

8. H. Wang, H. Owens, J. Smith, W. Chernicoff, M. Mazari, M. Pourhomayoun: "An End-to-End Traffic Vision and Counting System Using Computer Vision and Machine Learning: The Challenges in Real-Time Processing", Proceeding of Int. Conf. on Advances in Signal, Image & Video Processing, 2018 (Best Paper Award).

9. M. Pourhomayoun et. al, "Real-Time Big Data Analytics for Traffic Monitoring and Management for Pedestrian and Cyclist Safety", Proceeding of The Conf. on Big Data, Small Data, Linked Data and Open Data, 2019.

10. US Department of Transportation (USDOT) Newsletter April 2021, https://www.transportation.gov/sites/dot.gov/files/2021-04/UTC%20Newsletter_Quarterly_April_Q2_3-18.pdf.

11. TRB newsletter, http://www.trb.org/Publications/PubsTRBENewsletter.aspx.

12. M. Piccardi, "Background subtraction techniques: a review", IEEE Int. Conf. on Sys, Man and Cybernetics, 2004.

13. T. Bouwmans, F. El Baf, B. Vachon, "Background Modeling using Mixture of Gaussians for Foreground Detection – A Survey"". Recent Patents on Computer Science, 2008.

14. R. Girshick, "Fast R-CNN", IEEE International Conference on Computer Vision (ICCV), 2015.

15. S. Ren et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017.

16. J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", IEEE Conf. on Computer Vision and Pattern Recognition, 2016.

17. P. Zarchan, H. Musoff, "Fundamentals of Kalman Filtering: A Practical Approach", ISBN 978-1-56347-455-2.

18. Project Website, https://transweb.sjsu.edu/mctm/research/utc/Traffic-Monitoring-and-Management-Pedestrian-and-Cyclist-Safety-Using-Deep-Learning-and-Artificial-Intelligence

19. Project Video, www.youtube.com/watch?v=01U5E4E8uyE&ab_channel=MinetaTransportationInstitute

About the Author

Mohammad Pourhomayoun, PhD

Dr. Pourhomayoun is an Associate Professor of Computer Science at California State University Los Angeles (CSULA). He is the founder and director of the Artificial Intelligence & Data Science Research Lab at CSULA. Dr. Pourhomayoun's research interests focus on data science, artificial intelligence (AI), and machine learning for social good applications including traffic management and safety, and urban sustainability.

MTI FOUNDER

Hon. Norman Y. Mineta

MTI BOARD OF TRUSTEES

Directors

MTI MINETA TRANSPORTATION INSTITUTE