PRINCIPAL ACADEMIC TUTOR
Sarah Azimi, Dipartimento di Automatica e Informatica (DAUIN), Politecnico di Torino

ACADEMIC TUTORS
Vincenzo Randazzo, Dipartimento di Elettronica e Telecomunicazioni (DET), Politecnico di Torino
Miriam Pirra, Dipartimento di Ingegneria dell'Ambiente, del Territorio e delle Infrastrutture (DIATI), Politecnico di Torino
Luca Cassano, Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano

EXTERNAL INSTITUTION
Getapper Software Development Agency

TEAM MEMBERS

Teo Bucci
Mathematical Engineering
Politecnico di Milano

Dario Cavalli
Computer Science and Engineering
Politecnico di Milano

Gabriele Gabrieli
Management Engineering
Politecnico di Milano

Mario Mastrandrea
Computer Engineering
Politecnico di Torino

Eimy Nicole Ramirez Rugel
Architecture - Built Environment - Interiors
Politecnico di Milano

Francesco Antonio Ravazzani
Architectural and Urban Design
Politecnico di Milano

Giuseppe Stracquadanio
Computer Engineering
Politecnico di Torino

S2CITIES

Toward Smart and Safe Cities: Exploiting Surveillance Videos for Real-time Detection of "Signal for Help"

Executive Summary

In the wake of increasing incidents of violence against women, the '*Signal for Help*' (SFH) has surfaced as an invaluable silent beacon. Developed as a discrete hand gesture for women under duress to indicate their plight, its importance became even more pronounced during the pandemic when virtual communication surged. The S2CITIES project seeks to amplify the efficacy of this signal through technology.

Using cutting-edge AI-driven hand gesture recognition, S2CITIES continually monitors surveillance footage. Once the SFH is detected, an **alert mechanism** is activated, instantly notifying the necessary **authorities**. To augment this system and address potential inaccuracies, a dedicated mobile application was devised. This app serves as a critical tool for law enforcement, allowing swift review and response to these alerts.
Training the AI required a novel approach, given the absence of an existing dataset for the SFH. Through widespread appeals to university students, a unique dataset was amassed.

The implications of S2CITIES are vast. It propels surveillance from passive monitoring to proactive intervention. Preliminary market analysis have pinpointed public transportation sectors and municipalities as primary beneficiaries, given their extensive surveillance infrastructure and the promise of enhanced **safety for women**.

However, the path to widespread adoption is not without challenges. Concerns around privacy, particularly in the private sector, pose obstacles. Some local businesses fear potential revenue losses, while others grapple with legal implications surrounding **surveillance**.

In essence, S2CITIES is more than just a tech initiative. It's a testament to the power of **innovation**, acting as a bridge between silent pleas for help and tangible, life-saving interventions.

Key Words
Signal for Help, Artificial Intelligence, Hand Gesture Recognition, Mobile Application, Domestic Violence
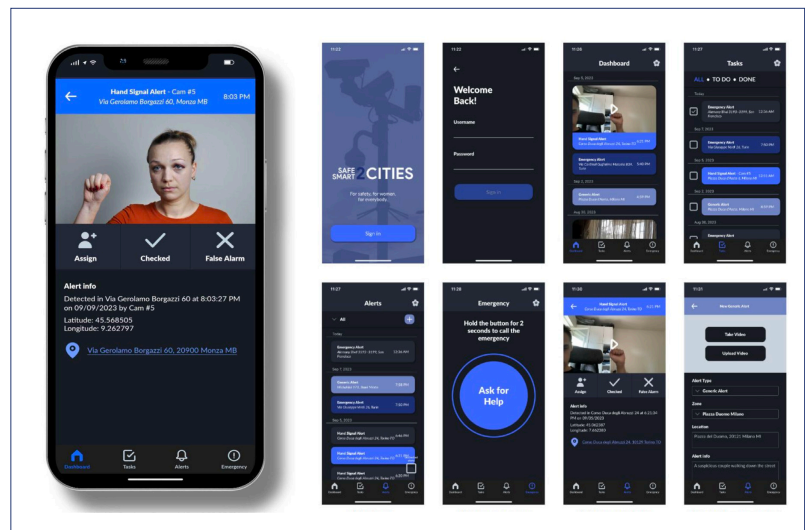
SAFE SMART 2 CITIES

Signal recorded by a **camera**

An in-app **notification** is sent to **authorities**

If the alert is verified as a **threat**,

Immediate assistance will be dispatched

1   2   3

For **safety**, for **women**, for **everybody**.

The intersection of technology and society has given rise to innovative solutions to societal issues. The tragic story of a 35-year-old woman in Rome serves as a stark reminder of the importance of using technology to help individuals in need. The woman was at dinner with her ex-partner when she felt unsafe and attempted to use a "signal for help". Unfortunately, nobody recognized the signal, and she was forced to leave the restaurant with her ex-partner, who later shot and killed her.

But what was the signal that the woman tried to use? The signal or the hand gestures is known as "**Signal for Help**". It is a hand signal that can be used to silently communicate to others that you need help and are in danger. The signal was first created by the Canadian Women's Foundation during the COVID-19 pandemic in response to increasing rates of **domestic violence**. Since then, it gained global recognition and has been used to provide a discreet way for victims of abuse to communicate their need for help, as the young woman in a restaurant in Rome tried to do.

This incident raises important questions about how technology can be used to prevent such tragedies from happening in the future. S2CITIES, with their advanced infrastructure and connectivity, present a unique opportunity to leverage technology to create safer communities. By adopting technologies such as IoT devices and AI-powered surveillance systems, we can create a system that can automatically recognize and respond to emergency signals, including the "Signal for Help."

In this project, we propose a system to recognize emergency signals and alert the appropriate authorities. By integrating surveillance cameras, and **AI algorithms**, the system can detect the "Signal for Help", and alert authorities in real-time. This system could be implemented in public spaces such as restaurants and malls, providing individuals with a discreet way to signal for help and increasing the chances of receiving assistance. This solution aims to enhance the effectiveness of the signal and increase the likelihood of individuals receiving the help they need.

Team description by skill

The S2CITIES team boasts a diverse range of experts from various domains, each of whom played a pivotal role in the project's success. Their varied **backgrounds** and unique contributions propelled the project forward, ensuring a well-rounded approach to the challenges.

Teo Bucci, a mathematical engineer, developed the robust dataset collection pipeline, as well as the final model exploiting MediaPipe's hand joint keypoints and, with Giuseppe, created a real-time inference framework. **Dario Cavalli**'s expertise in computer science powered the development of deep learning models, especially through the 3D-CNN approach, and forged academic collaborations. **Gabriele Gabrieli**'s finance background helped pinpoint key institutional partners, backed by valuable local interview data. **Mario Mastrandrea** spearheaded the creation of the S2CITIES Mobile app, melding it seamlessly with the gesture recognition system and crafting a versatile backend. **Eimy Nicole Ramirez Rugel,** with an architectural background, researched violence against women, surveillance landscapes, and project implementation sites. She also assisted in designing the mobile app's user experience. Despite remote challenges, **Francesco Antonio Ravazzani**'s urban design knowledge enhanced the machine learning aspects through video sorting and camera analysis. **Giuseppe Stracquadanio**'s computer engineering skills were pivotal, transforming theory into application and refining the 3D Convolutional Neural Network processes.

Goal

At the heart of the S2CITIES project lies a profound ambition: to re-envision urban environments as epitomes of safety and technological innovation. The team pursued this by conceptualizing an integrated system that promotes both the "**Smart**" and the "**Safe**" aspects of city living. In an era where technology is intertwined into nearly every facet of daily life, the project underscores the belief that cities can be both intelligent and protective, ensuring their citizens can thrive in secure, responsive environments. Particularly in the realm of **women's safety**, a domain often overlooked or inadequately addressed, the S2CITIES team set out with a clear vision to harness cutting-edge technological advancements to generate tangible, transformative impacts on **urban safety** dynamics.

Understanding the Problem

The urban tapestry is marred by persistent issues, and chief among them is the looming threat of **violence against women**. This menace, ever present in many city corners, necessitates more than just conventional solutions; it **demands innovation and foresight**. Existing surveillance mechanisms, while instrumental in providing a sense of security, often fall short in delivering real-time, actionable insights that could preempt potential threats or respond swiftly when they occur.
In the rapidly evolving landscape of smart cities, there is an exigent **call for solutions** that not only integrate effortlessly with urban infrastructures but are also sensitive to the varied and complex needs of its inhabitants.

Hierarchy of stakeholders
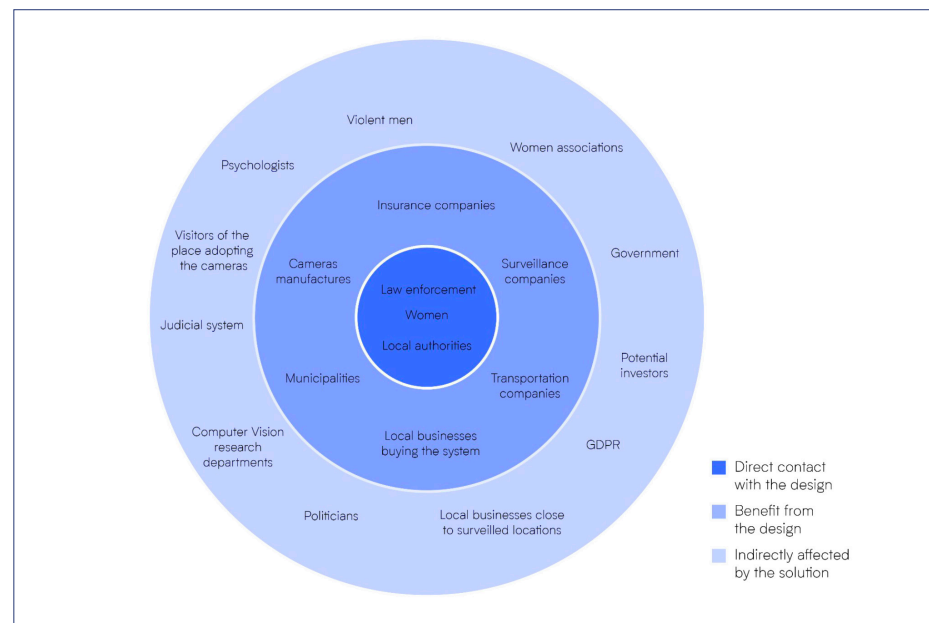
Exploring the Opportunities

The S2CITIES initiative delved deep into this realm of problems, sourcing invaluable information about the existing state of urban surveillance. By understanding the intricacies of violence against women in **public spaces**, the project was able to pinpoint specific urban areas that would benefit most from a transformative safety overhaul. This comprehensive understanding shaped the project's trajectory, ensuring its solutions were both grounded in reality and inspirationally forward-looking.

The growing global momentum around S2CITIES presented a fertile ground for innovation and collaboration. As cities and institutions grew increasingly cognizant of the value of these concepts, a tangible opportunity arose to embed them within urban planning and development frameworks. Ground-level interactions and dialogues with **local stakeholders** validated this enthusiasm, unveiling a genuine appetite for advanced safety solutions. Such positive feedback not only established a clear market potential but also anchored the project in a context of immediate relevance and utility.

Harnessing existing technological assets, such as advanced GPU clusters, gave the team a strategic advantage. This integration highlighted the possibilities of creating cutting-edge, yet efficient, **algorithms** that could operate seamlessly in **real-world scenarios**. These computational strengths, combined with the rich tapestry of skills and knowledge the team possessed, painted a promising picture. With expertise spanning the spectrum from the abstract precision of mathematical engineering to the human-centric focus of architecture and urban design, the project was equipped to navigate the intricate balance between technology and humanity. This ensemble of diverse thought processes and skills meant that the S2CITIES project was set to address challenges from a multi-lens, ensuring robust, relevant, and resonant **solutions** with the end-users' needs.

Figma board
with database
Architecture
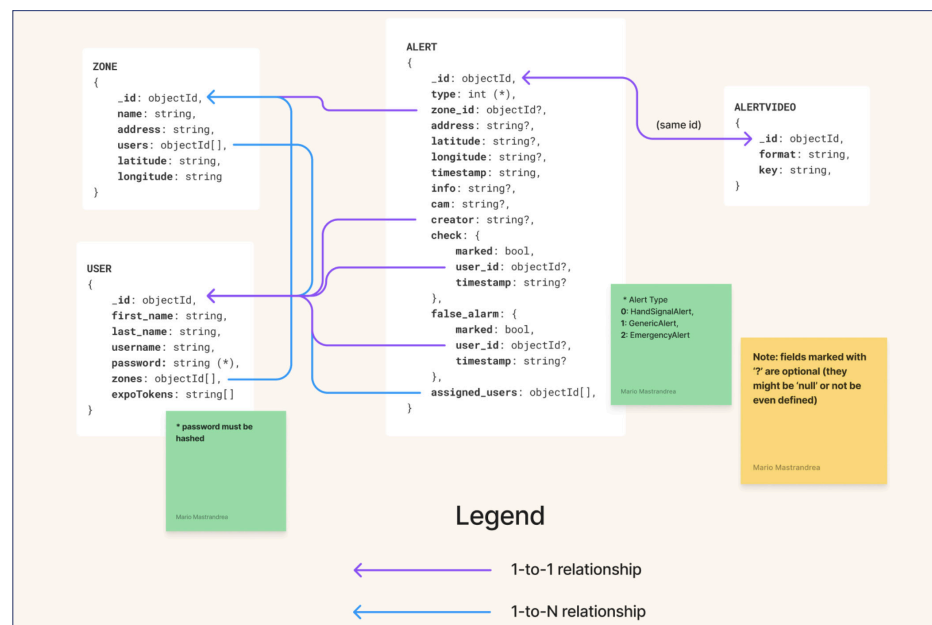
Mobile application
architecture
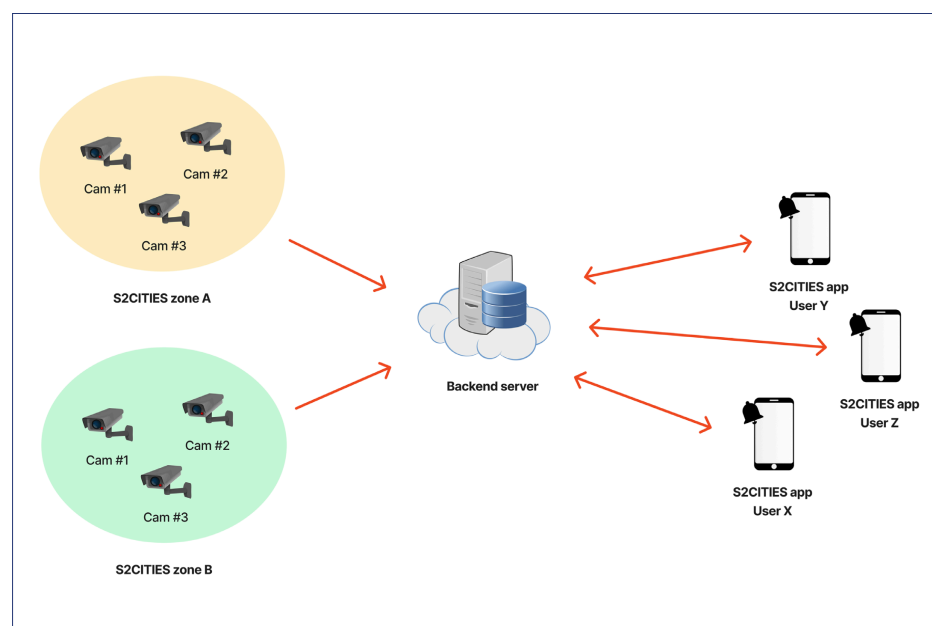
Generating a solution

The S2CITIES project represents the epitome of technological innovation and user-centric design, converging to form a holistic solution dedicated to enhancing urban safety. At the heart of the project lies a groundbreaking **hand gesture recognition system**, meticulously crafted to ensure utmost **accuracy and reliability**. Utilizing state-of-the-art tools such as MediaPipe, the project team has successfully established a robust foundation for precise hand gesture classification, marking a significant leap forward in this realm of technology.

The meticulous integration of this innovative technology into a user-friendly mobile application underscores the project's commitment to ensuring accessibility and engagement. Users can effortlessly navigate the diverse features of the app, further enriched by the seamless integration, thereby ensuring a satisfying and efficient user experience.
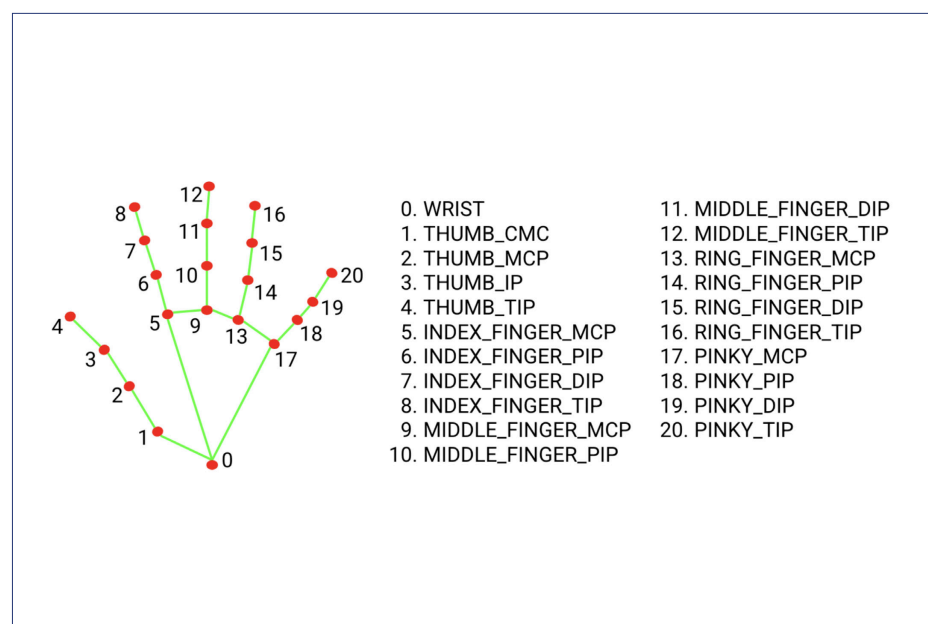
Despite initially founding the solution on custom-designed 3D-CNN architectures for video classification, the team recognized its limitations, specifically in recognizing the SFH gesture and efficiently classifying other gestures as negatives. This insight, combined with a comprehensive understanding of the potential pitfalls of analyzing features from entire videos, spurred a decisive shift in methodology.

The team adeptly harnessed an innovative approach, prioritizing the extraction of hand keypoints and converting them into a time series. This forward-thinking strategy eliminated the cumbersome need to process entire videos, swiftly discarding irrelevant information and markedly accelerating the entire operation. Subsequently, the time series classification was meticulously conducted using advanced feature extractors. Thousands of features were computed, with only the most significant retained through the use of sophisticated feature-selection methods. This deliberate and thoughtful process resulted in an **exceptionally lightweight model**, ensuring ultra-fast and efficient operation.
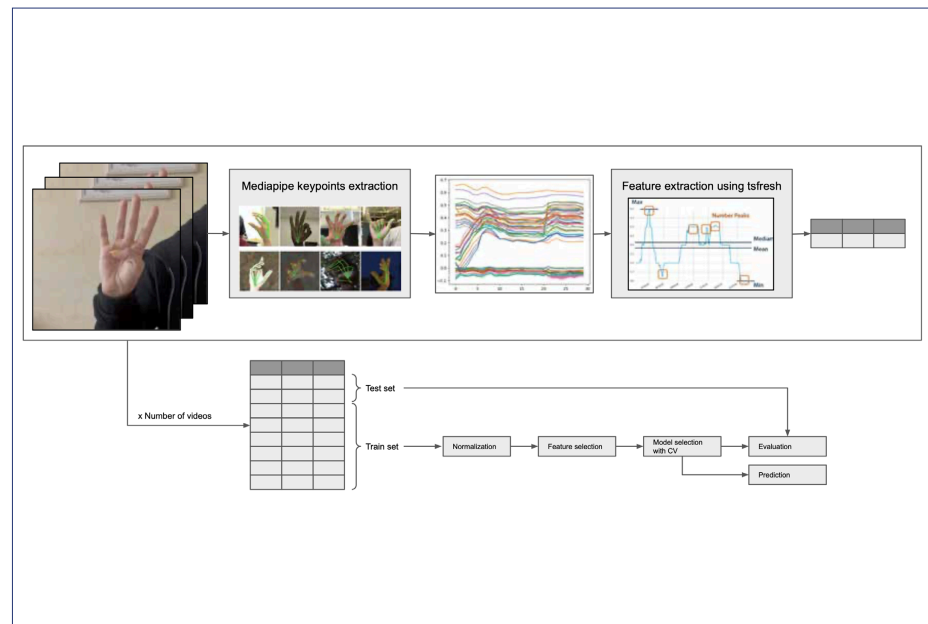
Abandoning the constraints of 3D-CNN architectures, the team seamlessly transitioned to a more **focused approach** centered on the extraction of hand key-points. This refined method guarantees a more targeted and efficient process for the final classification, systematically eliminating unnecessary data and substantially enhancing the accuracy and reliability of **gesture recognition**. This agile and adaptable approach underscores the team's commitment to continual improvement and innovation in the pursuit of excellence in gesture recognition technology.

Beyond the sophisticated technology and algorithms, the S2CITIES project showcases an unwavering commitment to design and **user experience**. The team's dedicated efforts have culminated in a solution that transcends technical excellence, embodying a visually alluring and user-centric design. This meticulous attention to detail ensures the creation of a harmonious **balance** between **functionality** and **aesthetics**, presenting users with a solution that is as compelling visually as it is technologically.

Full list of keypoints
extracted using
MediaPipe

0. WRIST	11. MIDDLE_FINGER_DIP
1. THUMB_CMC	12. MIDDLE_FINGER_TIP
2. THUMB_MCP	13. RING_FINGER_MCP
3. THUMB_IP	14. RING_FINGER_PIP
4. THUMB_TIP	15. RING_FINGER_DIP
5. INDEX_FINGER_MCP	16. RING_FINGER_TIP
6. INDEX_FINGER_PIP	17. PINKY_MCP
7. INDEX_FINGER_DIP	18. PINKY_PIP
8. INDEX_FINGER_TIP	19. PINKY_DIP
9. MIDDLE_FINGER_MCP	20. PINKY_TIP
10. MIDDLE_FINGER_PIP	

Final model
complete pipeline

In the strategic deployment of the project, particular emphasis is placed on optimizing its **impact** across various **public and private zones**. This deliberate and thoughtful placement ensures the maximization of safety enhancements, demonstrating the project's far-reaching implications for urban spaces.

In essence, the S2CITIES project emerges as a beacon of innovation, showcasing a seamless **amalgamation of advanced technology**, elegant design, and a keen understanding of urban safety needs. The resulting product stands as a resounding affirmation of the project's vision, offering an advanced, reliable, and aesthetically pleasing solution that holds the promise of significantly enhancing the **safety and security of urban environments**, thereby contributing to the creation of more secure, resilient, and sustainable cities for the future.

Main bibliographic references

- S. Azimi, C. De Sio, F. Carlucci and L. Sterpone. 'Fighting for a Future Free from Violence: A Framework for Real-Time Detection of "Signal for Help". In: Intelligent Systems with Applications 17 (February 2023), p. 200174. DOI: 10.1016/i. iswa.2022.2 00174.
- A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto and H. Adam. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. 16th April 2017. DoI: 10.48550/arXiv. 1704.04861.
- O. Köpüklü, A. Gunduz, N. Kose, Neslihan Kose, G. Rigoll, Gerhard Rigoll, Gerhard Rigoll and G. Rigoll. 'Real-Time Hand Gesture Detection and Classification Using Con-volutional Neural Networks'. In: 2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019) (14th May 2019), pp. 1-8. DoI: 10.1109/£g. 2019.8756576.
- C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg and M. Grundmann. 'MediaPipe: A Framework for Perceiving and Processing Reality'. In: 2019. DOI: 10.48 550/arXiv.1906.08172.
- J. Materzynska, G. Berger, I. Bax and R. Memisevic. 'The Jester Dataset: A Large-Scale Video Dataset of Human Gestures'. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Seoul, Korea (South): IEEE, October 2019, pp. 2874-2882. DOI: 10.1109/ICCVW.2019.00349.