Machine Learning-Based Radar Perception for Autonomous Vehicles Using Full Physics Simulation

ARIEN P. SLIGAR [iD]
ANSYS Inc., Canonsburg, PA 15317, USA
e-mail: arien.sligar@ansys.com

ABSTRACT Safety critical systems in Advanced Driver Assistance Systems (ADAS) depend on multiple sensors to perceive the environment in which they operate. Radar sensors provide many advantages and complementary capabilities to other available sensors but are not without their own shortcomings. Performance of radar perception algorithms still pose many challenges, one of which is in object detection and classification. In order to increase redundancy in ADAS, the ability for a radar system to detect and classify objects independent of other sensors is desirable. In this paper, an investigation of a machine learning based radar perception algorithm for object detection is implemented, along with a novel, automated workflow for generating large-scale virtual datasets used for training and testing. Physics-based electromagnetic simulation of a complex scattering environment is used to create the virtual dataset. Objects are classified and localized within Doppler-Range results using a single channel 77 GHz FMCW radar system. Utilizing a fully convolutional network, the radar perception model is trained and tested. The training is performed using a wide range of environments and traffic scenarios. Model inference is tested on completely new environments and traffic scenarios. These simulated radar returns are highly scalable and offer an efficient method for dataset generation. These virtual datasets facilitate a simple method of introducing variability in training data, corner case evaluation and root cause analysis, amongst other advantages.

INDEX TERMS ADAS, radar, FMCW, machine learning, millimeter wave, object detection, radar, YOLOv3.

I. INTRODUCTION

With autonomous vehicles on the horizon, performance requirements for Advanced Driver Assistance Systems (ADAS) are ever increasing. These systems will rely on multiple sensors to achieve either fully autonomous or semi-autonomous operation. Optical, radar, lidar and ultrasonic sensors are commonly used, either independently or in some form of sensor fusion to perceive the environment and traffic participants.

Perception algorithms, used to interpret and understand information from sensors, are critical in the operation of autonomous vehicles. Many avenues of research are being pursed in the development of these perception algorithms, with one major area of research being in the application of machine learning (ML). Optical systems using ML based

The associate editor coordinating the review of this manuscript and approving it for publication was Guolong Cui [iD].

perception algorithms are widely researched and already deployed in commercial systems [1]–[3]. Compared with optical systems, radar ML based perception algorithms are not nearly as mature. Recent research has demonstrated the feasibility of using ML to augment or replace traditional signal processing to interpret radar returns [4]–[10].

Modern radar sensors used in ADAS build a representation of the environment based on the observation of complex, scattered radio waves. From these scattered radio waves, information of an object's distance and velocity can be derived. The observation of these fields can be transformed into a Range-Doppler (RD) map as shown in Fig. 1. This result gives a visualization of all the scattered fields in terms of relative velocity and distance from the radar. These RD maps contain many ambiguous features that may not be easily identifiable to a human observer. Traditional radar signal processing, along with sensor fusion can be used to interpret these radar returns and perceive the environment. In this paper, an ML
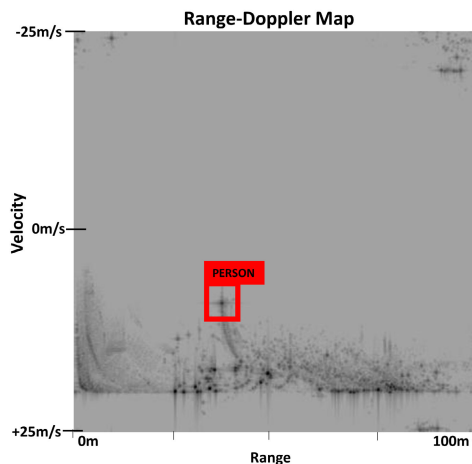
FIGURE 1. Example single channel Range-Doppler map output from the radar sensor, shown in colorized grayscale. Bounding box, and object classification label used in training overlay onto the plot, outlines the object of interest and location in Doppler/range.

FIGURE 2. Validation scenario: 2 corner reflectors.

based perception algorithm will be demonstrated that is used to classify and localize traffic participants within these RD maps.

Prior research by Zhang *et al.* [4] has shown that multi-channel radar returns can be used to train a convolutional neural network (CNN) to predict 3D object orientation of a single vehicle. Heuel and Rohling [5], [6], demonstrated the ability to use Doppler spectrum and range profiles to classify pedestrians using 24 GHz radar. Like optical image classification, the RD maps contain features that may be unique and can be recognized by a CNN. Research performed by [7]–[9] has shown how a CNN can be used to exploit features in the Doppler spectrum and range profiles to classify different objects. Work performed by [10] extends the classification to real world scenes by applying the CNN classification to regions of interest of the radar spectrum. These prior works demonstrate how the application of ML can be used in a radar perception algorithm. Furthermore, [11] demonstrated how physics-based simulations of optical sensors have been employed to train an ML model, which is then used for inference on physical/real camera sensors.

The approach demonstrated here, uses physics-based, simulated radar returns to train an ML based radar perception algorithm. Using a 77 GHz, Frequency-Modulated Continuous-Wave (FMCW) radar sensor, a perception algorithm will be trained to perform object detection and localization in real-world environments. Training will be done using only the RD map of a single channel radar. The perception algorithm will automatically discriminate between clutter, pedestrians and vehicles, not relying on any other traditional signal processing. Vehicles and pedestrians will be localized within the RD map using a bounding box defined by range (distance) and Doppler frequency (velocity). An example RD map used in the training and the corresponding label is shown in Fig. 1. A CNN is implemented and trained from scratch using only the simulated radar returns, in the form of RD
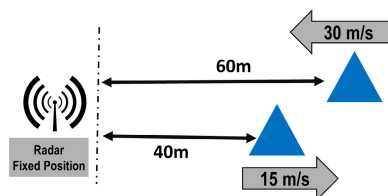
maps, for many realistic driving scenarios. When training is completed, the model inference is evaluated on new scenarios and environments never seen during the training process.

II. METHODOLOGY
A. PHYSICS BASED RADAR SIMULATION

Many experts believe hundreds of millions of miles need to be driven in order to demonstrate reliability [12]. This requirement will not likely be achieved using only real-world driven miles, but also rely on some form of simulation. From semantic simulations to physics-based simulations, there are many advantages offered when using simulations. Such as, scalability, implementation cost, scenario variability, crash replication, and sensor variation.

Training a robust ML radar perception algorithm is likely going to depend on having a large, varied dataset that represents the real radar returns of complex scenarios. As most real-world ADAS scenarios exceed 100k wavelengths in any given dimension at 77 GHz, a numerical technique needs to be selected that allows for accurate and efficient solution of problems at this scale. The shooting bouncing rays (SBR) technique is an asymptotic ray tracing approximation that can efficiently calculate scattered fields for electrically large geometries [13], [14]. This full physics simulation is based on the geometric and physical optics numerical techniques and can generate accurate electric and magnetic scattered fields in a multi-path environment. While capturing interaction of real material properties and complex geometry. An animated time-sequence of geometry is used to define the ADAS scenario simulation. Each time step in the series is analyzed by first calculating currents on all surfaces in the model using a user specified antenna pattern, and then radiating those currents back to another antenna pattern using the SBR method. The usage of the SBR technique has been demonstrated by [15] and [16] to perform full physics simulation of a realistic ADAS scenarios.

RD maps for a chirp sequence (CS) - FMCW radar can be achieved by simulating a sequence of chirps across a coherent processing interval (CPI). Each chirp is used to generate a time domain response, which can be scaled for distance. Fourier processing the chirps in a single CPI yields a Doppler frequency (velocity) response. The time dependent geometry describing the ADAS scenario is simulated for each chirp in the CPI using SBR. Generating scattered fields that represent the true position of all geometry at the corresponding time step and further processed into RD maps. The SBR solver implemented in ANSYS HFSS also includes a capability to
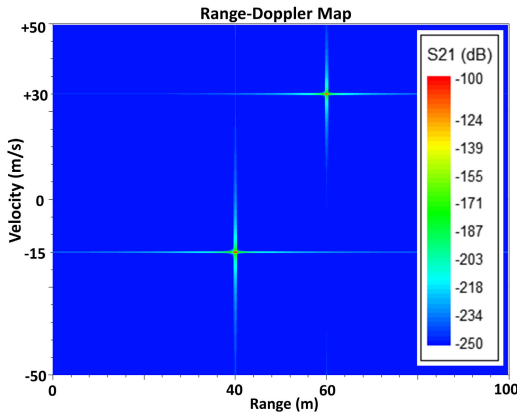
FIGURE 3. RD map from validation scenario simulation. Location of peak scattered fields agree with ground truth of (-15 m/s, 40 m) and (+30 m/s, 60 m) for each reflector. Peak values also agree with theoretical received power determined by radar range equation for an RCS value of 9 dBsm.

further accelerate the solution of CS-FMCW radar systems using a technique referred to as Accelerated Doppler Processing (ADP) [17]. ADP accelerates the solution by providing scattered fields for all chirps in one CPI in the same amount of time that would be required for simulation of a single chirp. This results in a simulation time that is typically several order of magnitude faster than if each individual chirp needed to be explicitly simulated.

A simple validation of the SBR and ADP simulation methodology is demonstrated by comparing simulated results with the ground truths and theoretical received power of a known scenario. The scenario to be simulated is outlined in Fig. 2, using two 9 dBsm corner reflectors. The radar module used for this validation example uses a single transmit and single receive antenna, each with peak total gain of 17.2 dB. The half-power beam width of each antenna is equal to 13.2 deg in the elevation plane and 51 deg in the azimuth plane. With a center frequency of 77 GHz, bandwidth = 600 MHz, CPI = 9.7 ms, and a pulse repetition frequency = 51.4 kHz issuing 500 chirps. The simulation generates the RD map shown in Fig. 3. The ground truth for velocity and range correspond with the peak locations seen in the simulated RD map. Theoretical received power is compared with simulated results by evaluating the radar range equation,

$$P_{received} = P_t G^2 \left(\frac{\lambda}{4\pi R} \right)^2 \frac{\sigma}{4\pi R^2}. \tag{1}$$

where P_t is transmitted power, G is transmit/receive antenna gain, σ is radar cross section of each target, λ is the wavelength, and R is range of the respective target. The simulated received power for each reflector is -101.05 dB and -108.1 dB. This agrees very closely with the theoretical received power of -101.4 dB and -108.5 dB.

B. TRAINING DATA GENERATION
Radar simulation of an ADAS scenario rely on a physical and electrical description of the environment, scenario and radar
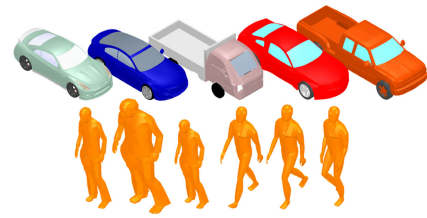

FIGURE 4. Library of actors, consisting of multiple vehicles and pedestrians.

performance characteristics. The environment consists of a physical geometry description, including terrain, roadway, signage, vegetation, buildings, etc., along with the electrical material property definitions. A scenario is defined by the placement of any traffic participants/actors into these environments, along with their respective position and trajectory. These actors also include all relevant electrical material properties and geometry features. The radar performance characteristics are defined by a far-field antenna pattern versus frequency for each transmitter and receiver used in the radar module.

A practical solution needs to be implemented to automate scenario generation at a scale required for training an ML model. The solution implemented here, begins with the creation of a library of environments and a library of actors. The library of actors includes 5 different vehicles and 6 different pedestrians (Fig. 4). The environment library consists of $\sim 200 \times 200$ meter section of the world with all objects and material properties defined. Material properties for the environment and traffic participants have been defined based on [18]–[20]. As an example, Environment 1 (Fig. 2 (a)) shows a two-lane road with a guard rail on one side ending in a tunnel and hillside. Environment 2 (Fig. 2 (b)) shows an intersection with several buildings, curbs and traffic lights. Along with the geometry, the environment also includes a description that specifies rules for a valid placement and trajectory of any actor. For instance, one such rule for an environment may describe that a vehicle actor can be placed anywhere on the roadway, but the main component of its velocity must be in the direction of travel of the respective lane. Multiple rules per environment are created to allow for wide variability in scenarios, and not limiting unique, unexpected scenarios, (i.e. a car traveling perpendicular to the road's primary direction of travel).

A scenario is evaluated for a single frame of data, where a frame is defined and evaluated over a single CPI. The scenario is generated by a random selection of up to 5 vehicles and/or pedestrians in a randomly selected environment. The location of the actors, along with their trajectories, are chosen randomly, but with respect to the environment rules. This process allows large-scale scenario generation to be automated, only limited by the manual process of generating the initial environment and rules.

Using the generated scenarios, simulation of the electrical radar characteristics is completed using the

HFSS-SBR+ solver. The radar system implemented here has a range resolution of 25 cm and maximum range ambiguity of 100 m. The velocity ambiguity is $+/-$ 25 m/s and velocity resolution of 0.12 m/s. This radar performance is achieved by issuing 416 up-chirps over a CPI equal to 16.2 ms with a Pulse Repetition Frequency of 25.7 kHz. Each chirp (sampled in the frequency domain) is composed of 416 samples over a 625 MHz bandwidth. The single transmit and single receive antenna configuration described in the simple validation case is also implemented. These parameters are typical of a generic medium range radar module used in automotive applications [21]. The perception algorithm performance will likely be impacted by these radar parameters and could be explored in future work.

Actor movements are restricted to be purely translational, without including movements that would contribute to what would be considered the micro-Doppler (MD) signature. For example, vehicles are translated in position for the defined velocity, but with no rotation in the wheels. Similarly, pedestrians are translated at a single velocity, with extremities held in a fixed pose. The translation of actors is captured by the SBR solver by positioning all geometry in the environment to reflect the current location dependent on the time step requested. The animated sequence would appear as cars and pedestrians "sliding" in position versus time. At each time step a full physics extraction of the scattered fields is performed on the current state of all geometry. Although this technique will not capture these MD effects, the results will still capture the Doppler velocity spread due to an objects angular span relative to the radar. As well as any multi-body, and multi-path interaction of fields. Including MD signatures would likely generate additional features in the velocity dimension of the RD map, ultimately improving detection performance [10]. MD effects were excluded to reduce computational complexity of the simulations in this initial study. With the completion of each scenario, the resulting RD map is exported and stored along with all corresponding labels.

Accurate annotation/labeling of radar returns needed for ML training can be challenging in a physical measurement scenario, but in simulation the labels are available with virtually no additional effort or cost. This ground truth data is inherently available and automatically synced with the output data in simulation.

As all scenarios are completely independent of one another, embarrassingly parallel distributed computing can be leveraged to scale simulation results. For the given scenarios, using a 128-core compute cluster, simulation results can be generated at a rate of approximately 500 scenarios per hour.

C. MODEL ARCHITECTURE
The radar perception algorithm demonstrated will leverage many of the advances in image-based ML object detection that have occurred in recent years. Namely, the model architecture implemented here, is based on the YOLOv3 architecture [22]. The YOLOv3 object detection architecture is widely used in many optical systems, with advantages being

accuracy and speed of inference. These advantages make it a good candidate for the time critical detection systems needed in autonomous vehicles.

Simulation results for each scenario are transformed into a colorized image of Doppler vs. range by taking the magnitude of the complex scattered field quantities. This image, also referred to as a RD map (Fig. 1), is the input to the training. Labels for the corresponding image are generated during the simulation setup.

Pre-processing of the radar image is limited to a thresholding of -80 dB from the peak detected values (per scenario). Opportunities for image augmentation, and other image pre-processing may exist to improve detection and training results but will not be explored at this time.

III. RADAR OBJECT DETECTION
A. TRAINING
The CNN based on the YOLOv3 architecture was trained from scratch, using simulation results generated as described in section II B. Training was completed over 25 epochs using 9000 scenarios, and validated on 900 scenarios. Individual images are generated as a single frame, with no reference to the previous timestep/frame in the scenario. A future opportunity to explore tracking and/or CNN-LSTM (Long Short-Term Memory) networks [7] is possible to exploit the transient nature of real-world scenarios.

B. TESTING AND RESULTS
Testing was performed on two different datasets, with the model performance evaluated individually on each dataset. The first dataset consists of ∼900 scenarios using the same environments/terrain observed in the training, but with unique scenarios not seen by the original training. The second dataset is evaluated on 900 unique scenarios simulated in a completely new environment never seen by the training. The traffic participant library is kept consistent between both datasets in testing and training. This new environment in the second dataset demonstrates the robustness of the inference model. Model performance was evaluated using mean Average Precision (mAP) [23], and performance for each class evaluated by Average Precision (AP).

1) TEST DATA SET – 900 NEW SCENARIOS
Evaluating precision across all 900 scenarios (Fig. 6), an 83.8% mAP was achieved. The model inference AP for vehicles is 94% and 74% for pedestrian. A higher range resolution radar would likely improve performance on pedestrian detection. The pedestrians occupy far fewer range bins than vehicles, resulting in fewer features in the range profile. Also, as previously mentioned, simulated results do not include the MD signatures. This is likely the primary means of differentiating pedestrians from the many similar sized, stationary targets, where MD would create unique features in both range and velocity that could be used in object detection.
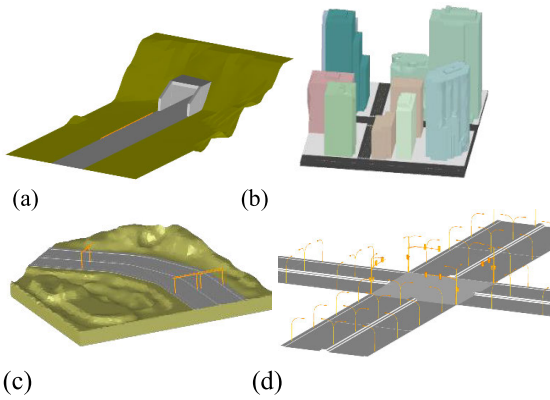
FIGURE 5. Library of environments. (a) Tunnel with hillside and guardrail. (b) City/urban intersection. (c) Curved roadway with signage. (d) Intersection with traffic signals and streetlights.
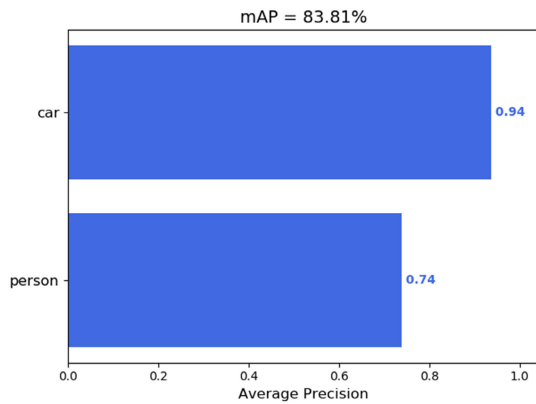

FIGURE 6. Testing results for dataset using 900 new scenarios in previously seen enviroments. Evaluation of mAP and AP of the test cases.
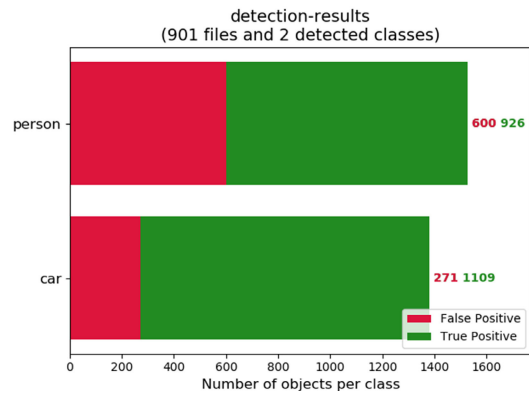

FIGURE 7. Testing results for dataset using 900 new scenarios in previously seen enviroments. Detection results showing false/true positives for each class.

An example of inference on a single scenario, never seen by the original training, is shown in Fig. 8. Object localization and classification is predicted using only the RD map as an input to the trained model. For a qualitative visualization of the inference, the ground truth is overlayed onto the RD map, along with the predicted object detections. Note the background clutter due to the environment was correctly ignored,
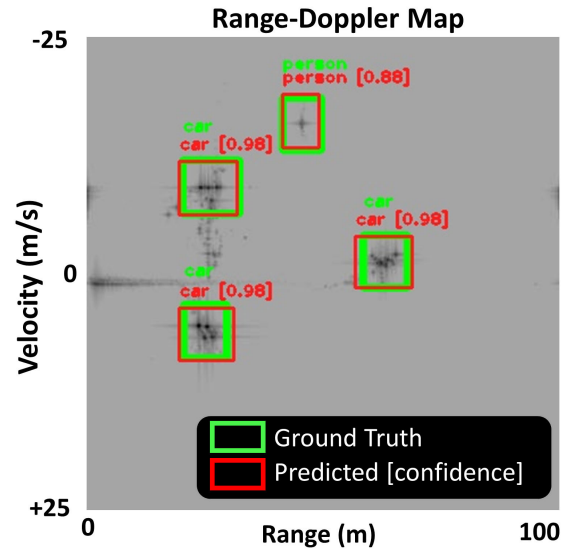

FIGURE 8. Example of inference on a single scenario. Red is predicted object detections based on trained model. Object detections shown in green are the ground truth overlay for purpose of visual comparison.
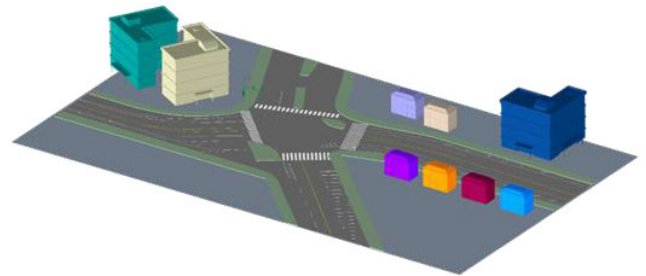

FIGURE 9. New environment used in testing, never seen by original training. The environment is defined by a complex intersection with buildings, traffic control devices (i.e. lights, cones, curbs, elevated islands).

and correct classification/localization of traffic participants was achieved.

2) TEST DATA –NEW SCENARIO AND NEW ENVIRONMENT

The new environment used to test model performance in a completely new context is shown in Fig. 8. Scenarios using this new environment are generated in a similar manner as described in section II. B. The performance of this model maintained a relatively high AP for vehicle detection but is reduced for pedestrians. This results in a decreased mAP from the previous dataset from 84% to 72% (Fig. 10). Figure 10 demonstrates model inference on one scenario using this new environment. Background clutter results in a higher percentage of false positives, as the model has difficulty in distinguishing between stationary objects and pedestrians. An example of a false negative is shown in Fig. 13. As previously discussed, this is likely to be improved by increased range resolution and incorporating MD signatures.

The performance of the previous dataset may have had the advantage of learned background clutter, resulting in
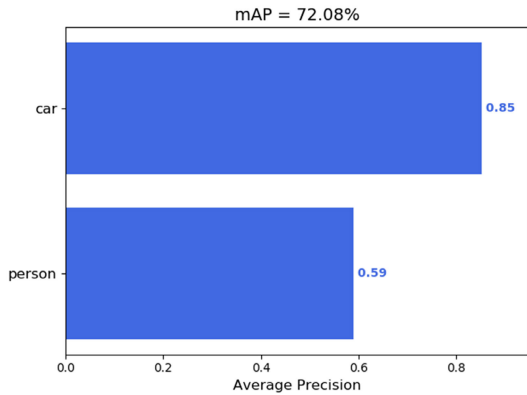
FIGURE 10. Results from testing model inference on completely new enviroment and scenarios. Evaluation of mAP and AP of the test cases.
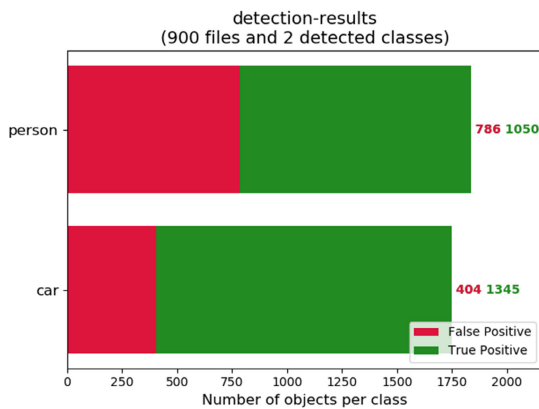

FIGURE 11. Results from testing model inference on completely new enviroment and scenarios. Detection results showing false/true positives for each class.
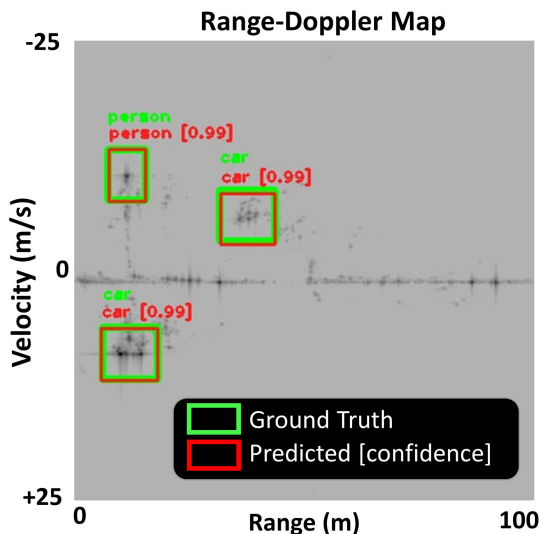

FIGURE 12. Results from a single scenario using the new environment, showing correct object detections.
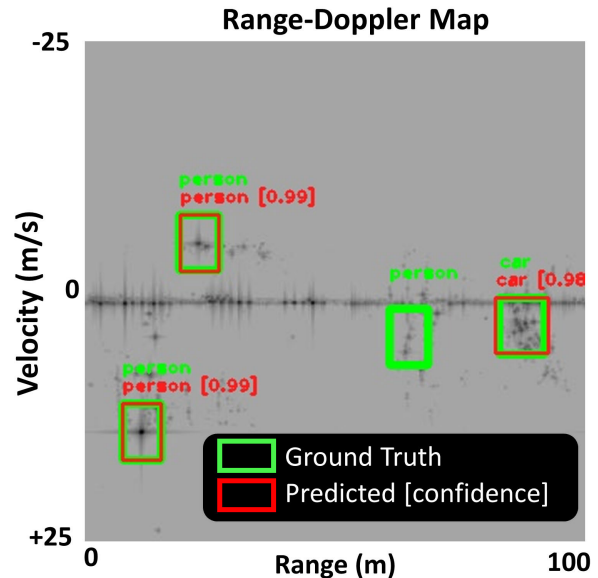
FIGURE 13. Results from a single scenario using the new environment, showing a pedestrian not correctly located or classified.

to introduce unique test data, explore the robustness of our inference model and easily investigate root cause failures.

IV. CONCLUSION

This paper demonstrates an application of ML based perception algorithm for radar systems. Without employing traditional signal processing and relying solely on an image-based object detection model, targets were classified, and localized in range-Doppler. The ML model training was based on simulated, physics-based radar returns. Simulation allows for a wide variability in scenarios and environments, all while providing automatically annotated output. Physics based simulation can prove to be a critical tool in development perception algorithms required for ADAS.

REFERENCES

[1] A. Moujahid, M. ElAraki Tantaoui, M. D. Hina, A. Soukane, A. Ortalda, A. ElKhadimi, and A. Ramdane-Cherif, "Machine learning techniques in ADAS: A review," in *Proc. Int. Conf. Adv. Comput. Commun. Eng. (ICACCE)*, Paris, France, Jun. 2018, pp. 235–242, doi: 10.1109/ICACCE.2018.8441758.

[2] Tesla. (2020). *Autopilot*. [Online]. Available: https://www.tesla.com/autopilot

[3] Mobileye. (2020). *Technology*. [Online]. Available: https://www.mobileye.com/our-technology/

[4] G. Zhang, H. Li, and F. Wenger, "Object detection and 3D estimation via an FMCW radar using a fully convolutional network," *Comput. Res. Repository*, to be published, doi: 10.13140/RG.2.2.15796.91521.

[5] S. Heuel and H. Rohling, "Pedestrian classification in automotive radar systems," in *Proc. 13th Int. Radar Symp.*, Leipzig, Germany, May 2012, pp. 477–484.

[6] S. Heuel and H. Rohling, "Pedestrian recognition based on 24 GHz radar sensors," *Proc. 11th Int. Radar Symp.*, Vilnius, Lithuania, Jun. 2010, pp. 1–6.

[7] A. Angelov, A. Robertson, R. Murray-Smith, and F. Fioranelli, "Practical classification of different moving targets using automotive radar and deep neural networks," *IET Radar, Sonar Navigat.*, vol. 12, no. 10, pp. 1082–1089, Oct. 2018, doi: 10.1049/iet-rsn.2018.0103.

better performance over the completely new environment. This new environment not only demonstrates the limitation of this model, but also reveals how simulation can be used

[8] E. A. Hadhrami, M. A. Mufti, B. Taha, and N. Werghi, "Ground moving radar targets classification based on spectrogram images using convolutional neural networks," in *Proc. 19th Int. Radar Symp. (IRS)*, Bonn, Germany, Jun. 2018, pp. 1–9, doi: 10.23919/IRS.2018.8447897.

[9] S. Capobianco, L. Facheris, F. Cuccoli, and S. Marinai, "Vehicle classification based on convolutional networks applied to FM-CW radar signals," 2017, *arXiv:1710.05718*. [Online]. Available: http://arxiv.org/abs/1710.05718

[10] K. Patel, K. Rambach, T. Visentin, D. Rusev, M. Pfeiffer, and B. Yang, "Deep learning-based object classification on automotive radar spectra," in *Proc. IEEE Radar Conf. (RadarConf)*, Boston, MA, USA, Apr. 2019, pp. 1–6, doi: 10.1109/RADAR.2019.8835775.

[11] S.-A. Schneider and K. Saad, "Camera behavioral model and testbed setups for image-based ADAS functions," *Elektrotechnik Informationstechnik*, vol. 135, nos. 4–5, pp. 328–334, Jul. 2018, doi: 10.1007/s00502-018-0622-7.

[12] H. Winner, S. Hakuli, F. Lotz, and C. Singer. *Handbook of Driver Assistance Systems*. Accessed: Dec. 1, 2019. [Online]. Available: https://www.rand.org/content/dam/rand/pubs/research_reports/RR1400/RR1478/RAND_RR1478.pdf

[13] H. Ling, R.-C. Chou, and S.-W. Lee, "Shooting and bouncing rays: Calculating the RCS of an arbitrarily shaped cavity," *IEEE Trans. Antennas Propag.*, vol. 37, no. 2, pp. 194–205, Feb. 1989, doi: 10.1109/8.18706.

[14] R. A. Kipp, "Curved surface scattering geometry in the shooting and bouncing rays method," in *Proc. IEEE Antennas Propag. Soc. Int. Symp.*, Toronto, ON, USA, Jul. 2010, pp. 1–4, doi: 10.1109/APS.2010.5562313.

[15] U. Chipengo, "Full physics simulation study of guardrail radar-returns for 77 GHz automotive radar systems," *IEEE Access*, vol. 6, pp. 70053–70060, 2018.

[16] J. D. Castro, S. Singh, A. Arora, S. Louie, and D. Senic, "Enabling safe autonomous vehicles by advanced mm-wave radar simulations," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Boston, MA, USA, Jun. 2019, pp. 1476–1479.

[17] *Ansys*, Canonsburg, PA, USA, 2019, pp. 1695–1697.

[18] S. Gabriel, R. W. Lau, and C. Gabriel, "The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues," *Phys. Med. Biol.*, vol. 41, no. 11, pp. 2271–2293, Jan. 1999.

[19] S. S. Zhekov, O. Franek, and G. F. Pedersen, "Dielectric properties of common building materials for ultrawideband propagation studies [Measurements Corner]," *IEEE Antennas Propag. Mag.*, vol. 62, no. 1, pp. 72–81, Feb. 2020.

[20] E. S. Li and K. Sarabandi, "Low grazing incidence millimeter-wave scattering models and measurements for various road surfaces," *IEEE Trans. Antennas Propag.*, vol. 47, no. 5, pp. 851–861, May 1999.

[21] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 845–860, Mar. 2012.

[22] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*. [Online]. Available: https://arxiv.org/abs/1804.02767

[23] *Detection Evaluation*. Accessed: Dec. 1, 2019. [Online]. Available: http://cocodataset.org/#detection-eval

ARIEN P. SLIGAR received the B.S. and M.Sc. degrees in electrical engineering from Oregon State University, in 2004 and 2006, respectively, with a focus on electromagnetics and microwave components.

In 2006, he joined ANSYS Inc., where he is currently a Principal Engineer, focusing on advanced applications of numerical simulation for electromagnetics and electronics. He is an expert in the application of electromagnetic field simulation to the design of antennas, complex antenna systems, microwave components, and high-speed electronics. He works with leading technology companies and provides engineering guidance allowing them to successfully apply simulation and automated workflows to their most difficult design challenges.

● ● ●