IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 11, 2021, accepted January 20, 2021, date of publication January 25, 2021, date of current version February 1, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3054368

Complex-Valued Channel Attention and Application in Ego-Velocity Estimation With Automotive Radar

HYUN-WOONG CHO, SUNGDO CHOI, YOUNG-RAE CHO, AND JONGSEOK KIM
Samsung Advanced Institute of Technology, Samsung Electronics, Suwon 16678, South Korea
Corresponding author: Hyun-Woong Cho (hwoong.cho@samsung.com)

ABSTRACT Attention mechanisms have been widely integrated with various neural networks to boost performance. However, when an attention mechanism was applied to a radar ego-velocity estimation network, the importance of carefully handling the amplitude and phase of complex-valued tensor was revealed. Therefore, in this study, we present a self-attention mechanism designed to handle complex-valued tensors in order to capture the rich contextual relationships implied within amplitude and phase. To exploit the advantages of complex-valued attention ($\mathbb{C}$A), we evaluated its impact while performing ego-velocity estimation tasks based on radar data, whose amplitude and phase are related to the electromagnetic scattering of the target being observed. Radars are suitable sensors for such tasks as they are capable of long-range detection and instantaneous velocity measurement under variable weather and lighting conditions. In particular, we coupled our $\mathbb{C}$A module with complex-valued neural networks, known to be particularly powerful for handling wave phenomena. The proposed method exhibits robust estimation performance, regardless of whether the Doppler ambiguity problem occurs and eliminates the dependence on preprocessing stages, including target detection and static target indication. Furthermore, it achieves improved stability during training via geometrical constraint regularization, and implicitly allows velocity conversion between the sensor and vehicle frames, even if the mount information of the sensor was not provided. Finally, ablation experiments conducted on extensive real-world datasets show noticeable improvement in estimation performance.

INDEX TERMS Automotive radar, complex-valued attention, complex-valued neural network, deep learning, ego-velocity estimation.

I. INTRODUCTION

Automotive radars have gained significant attention in recent years and now have applications in numerous fields, including advanced driver assistance systems and self-driving cars. In particular, 79-GHz millimeter-wave (mmWave) radars have become essential in commercial vehicles because of their high performance, high integration, small volume, and low cost [1]. Further, the latest generation of automotive radars has opened up new possibilities for advanced algorithms.

The knowledge of ego-motion is an indispensable presupposition associated with advanced applications, such as grid-mapping, state feedback for vehicle control, route planning and localization, and tracking objects in their surroundings [2]. Unlike vision and LiDAR systems, radars can measure relative motion instantaneously, even under variable lighting and weather conditions. Furthermore, radar receives data-efficient and information-rich signals, which include high-accuracy location, velocity, and angle estimates of objects. Because of these practical benefits including instantaneous Doppler velocity measurement, radars are suitable for the motion estimation of the ego-platform.

Deep learning has emerged as a powerful technique deployed in diverse fields and has recently been applied to automotive radar systems in the fields of target detection [3], tracking [4], classification [5], [6] and activity recognition [7]. Radar data are not acquired as an image, instead they are inherently complex-valued tensors, whose amplitude and phase are related to the electromagnetic (EM) scattering of the target being observed. However, most previous approaches that have applied convolution neural networks

The associate editor coordinating the review of this manuscript and approving it for publication was Wenjie Feng.

to radar data, discard phase information by summing the signal power over a specific dimension [3], preprocessing the measured point cloud into a grid [4], accumulating data using the occupancy grid technique [5], applying non-coherent integration on the range-velocity spectrum [6], or converting the spectrogram to gray scale [7]. Because the phase of radar data contains abundant information, a delicate network design is required to completely incorporate radar information. In our previous study, we proposed a radar ego-velocity estimation network (REVEN) [8], which was structured using complex-valued building blocks to address the characteristics of complex-valued radar data.

Attention mechanisms have been widely integrated with various neural networks to boost performance based on the simple, yet powerful, premise that we attend to a certain part to process a considerable amount of information. However, when we applied an attention mechanism to REVEN, the importance of carefully handling the amplitude and phase of complex-valued tensor was revealed.

In this study, we introduce a novel complex-valued attention module to completely leverage the abundant information of radar data. To the best of our knowledge, this is the first study on an attention mechanism designed to handle 3D complex-valued tensors. Our complex-valued attention mechanism is incorporated into REVEN and evaluated using an extensive radar dataset acquired during driving through the city for an ego-velocity estimation (EVE) task. Our main contributions can be summarized as follows:

- An end-to-end (E2E) complex-valued neural network architecture is proposed that directly handles raw radar data, thus overcoming the preprocessing dependence and Doppler ambiguity problems.
- A complex-valued attention mechanism is designed to completely incorporate information in a complex-valued tensor.
- The estimation performance is enhanced using geometric relations as a training constraint.
- The performance is evaluated using extensive datasets comprising over 100,000 radar frames.

The remainder of this paper is organized as follows. We review the related literature, covering ego-motion estimation, self-attention, and speech enhancement in Section II. In Section III, we describe our framework by introducing the 3D complex-valued tensors of the radar signal, presenting the proposed complex-valued attention ($\mathbb{C}$A) module, addressing the overall network architecture used for EVE, and discussing the geometric constraints for the network training. The database configuration and simulation results are reported in Section IV. The paper is concluded in Section V.

II. RELATED WORK

Radars are mounted onto the vehicle's chassis with a lever-arm offset from the reference frame of the ego-vehicle. This offset provides the foundation for estimating the rotational motion based on the observed relative radial motion [2], [9]–[14]. In a seminal work [9], the ego-motion was estimated based on the sinusoidal relation between the measured Doppler velocities and the azimuth angles. Furthermore, a random sample consensus algorithm [15] was used to discriminate the static targets for which the sinusoidal relation was established. This work, which determined an ego-velocity vector of 2 degrees of freedom (DoF), was extended to the case of multiple distributed radars that deals with the full 2D vehicle motion state, i.e., 3 DoF [10]. A probabilistic approach incorporating spatial registrations of radar scans was also proposed [11]. This joint spatial and Doppler-based estimation functions without lever-arm offsets or motion assumptions but involves significant computational costs. In subsequent research [12], the normal distribution transform algorithm [16] was utilized for faster spatial alignment, and the complexity was further reduced by deriving a sparse probabilistic representation [13]. A hybrid approach [14] was proposed to decouple translational and rotational motion by combining the benefits from scan matching and instantaneous approaches. However, because the underlying principles of these methods are based on well-discriminated static target information, their estimated performance highly depends on the preceding target detection stage. Therefore, performance may be limited in highly dynamic scenes or scenes in which few targets are detected.

Since attention mechanism was first introduced in natural language processing literature [17], [18], research has been actively conducted to expand its application to a wide range of fields. Accordingly, the attention module has been increasingly applied to vision literature. In the image generation task, the attention module was introduced to allow attention-driven long-range dependency modeling for generators [19]. Moreover, the effectiveness of formalizing self-attention as a non-local operation to model the spatio-temporal dependencies in image sequences was explored [20]. To design a deeper network structure, residual channel attention network (RCAN) [21] was used to improve the super-resolution performance by considering feature correlations in the channel dimension. Despite this progress, self-attention has not yet been explored in the context of complex-valued neural networks (CVNNs). Unlike previous works, this study aims to extend the self-attention mechanism to the automotive radar field and to carefully design an attention module that can capture rich contextual relationships implied within 3D complex-valued tensors.

In speech enhancement literature [22]–[25], which is another area that handles complex-valued data called spectrograms generated through short-time Fourier transform, multiplicative masks are known to perform better than alternative techniques, such as direct prediction of spectrograms or waveforms. The complex-valued phase has often been neglected because of the difficulty in its estimation. Nevertheless, the phase-sensitive mask (PSM), the first mask-based attempt, was proposed for incorporating phase information [22]. To address the performance limitation of PSM caused by reusing the noisy phase, complex-valued ratio mask (cRM) based approaches [23], [24] were proposed to

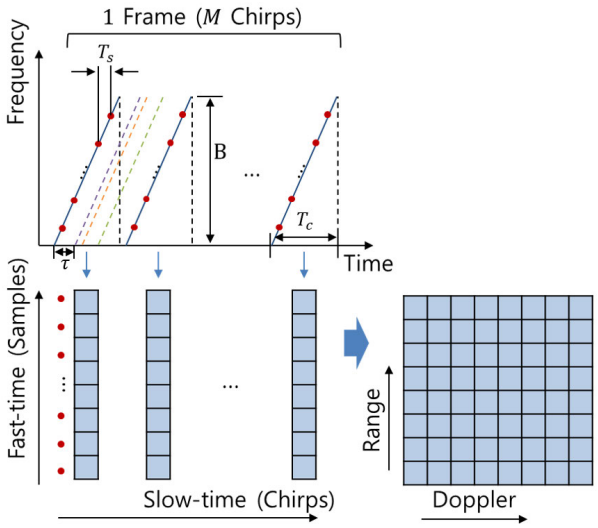

FIGURE 1. Illustration of the range-Doppler matrix formation for one-frame data of a fast-chirp FMCW radar.

directly optimize complex values. Recently, a deep complex U-Net [25] was proposed; it is a variant of U-Net [26] that combines the advantages of the cRM framework and a deep CVNN. Inspired by these masking approaches proposed to address the phase information of complex-valued data, the proposed $\mathbb{C}$A module was developed.

III. METHODS

At a high level, our $\mathbb{C}$A module is integrated with the architecture of REVEN and evaluated based on data collected using the frequency modulated continuous wave (FMCW) multiple-input multiple-output (MIMO) technology, which is widely adopted by most commercial radars. Because the angle-velocity relationship of static objects is directly related to the ego-velocity information, the radar data are rearranged so that the angle and velocity can be obtained from the spatial domain, and the range information is placed along the channel axis. Because attention mechanisms can select features, $\mathbb{C}A$ is applied in the form of a complex-valued channel attention ($\mathbb{C}$CA); this channel attention mechanism assigns large weights to channels that show a high response to salient objects to perform target detection in the range profile.

In this section, we detail the proposed approach, starting with the signal model, followed by the $\mathbb{C}A$ module and the network architecture. Finally, we introduce a new geometrical constraint (GC), which is important for the stable convergence of training and performance improvement.

A. SIGNAL MODEL

FMCW radars are widely employed in the automotive field owing to their advantages, such as light weight, low power consumption, and cost-effectiveness, while achieving relatively high range resolution. In the FMCW radar system, the frequency synthesizer generates a linear frequency ramp with a slope

$$\mu = \frac{B}{T_c} \tag{1}$$

where B is the bandwidth and T_c is the sweep time. The received signal is down-converted by the mixer and stored in a matrix after being sampled by an analog-digital converter with a sampling frequency of f_s. A beat frequency, f_b, of the resulting baseband signal includes the pure target range contribution, f_R, and the radial-velocity-dependent Doppler frequency, f_D, as follows:

$$\begin{aligned} f_b &= f_R - f_D \\ &= \mu\frac{2R}{c} + \frac{2}{\lambda}v \end{aligned} \tag{2}$$

where c is the speed of light, λ is the wavelength, R is the range between the target and radar, and v is the radial velocity of the target. Generally, the chirp duration, T_c, is sufficiently short for the range-dependent frequency, f_R, to predominate the beat frequency, i.e., $f_b \approx f_R$.

A sequence of M frequency ramps is transmitted during the one-frame interval, T_f, to resolve the range-velocity ambiguity given in (2). Given that each chirp has N sampling points, the M measurements with the same ramp are concatenated and form an $N \times M$ 2D array, i.e., a range-Doppler map, as shown in Fig. 1. A fast Fourier transform (FFT) along the slow-time directly yields the Doppler frequency, f_D.

In cells where the targets exist in the range-Doppler map, the phase difference between the antenna elements form a vector that is used for direction of arrival (DoA) estimation. Considering impinging signals are planar (far-field condition), a wave arriving at the k-th antenna element of the uniform linear array propagates an additional distance of $kd\sin\theta$, where d is the distance between adjacent antenna elements, as shown in Fig. 2.

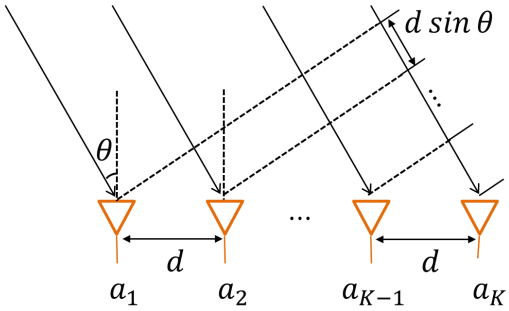

FIGURE 2. Geometry of uniform linear array.

The ADC sampled baseband signal comprises three dimensions with indices n, m, and k corresponding to the fast-time, slow-time, and antenna elements, respectively. The resulting signal can be expressed as follows:

$$\begin{aligned} \mathbf{s}(n,m,k) = A\exp\Bigg[i2\pi\Bigg\{&\left(\frac{2f_c v}{c} + \frac{2BR}{cT_c}\right)T_s n \\ &+ \frac{2f_c R}{c} + mf_D T_f + \frac{f_c kd\sin\theta}{c}\Bigg\}\Bigg] \end{aligned} \tag{3}$$

where f_c is the center frequency of the chirp, T_s is the sampling interval, A is the signal amplitude that is changed through propagation and reflection, and k is the position of virtual antenna elements, which varies depending on the MIMO configuration with the TX and RX index.

The range, velocity, and angle information is encoded along each dimension of the 3D array, as in expressed by (3). Instead of directly feeding the data into the network, we apply FFT on each dimension, which is a typical and lossless method to interpret FMCW data. The 3D array data after applying FFT can be expressed as

$$\mathbf{S}(p,q,r) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{k=0}^{K-1} s_w(n,m,k) \times \exp\left[-i2\pi\left(\frac{np}{N_R}+\frac{mq}{N_V}+\frac{kr}{N_A}\right)\right] \quad (4)$$

where $s_w(n,m,k)$ is $s(n,m,k)$ after an appropriate windowing operation is applied; K is the number of virtual array elements; and N_R, N_V, and N_A are the numbers of FFT points for each dimension. This preprocess reduces the capacity of the network needed, and can be selectively applied to each dimension.

Although it is common practice to treat data at each dimension separately, data between dimensions are not perfectly separable in practice [27]. Thus, the use of sub-space type algorithms, such as multiple signal classification [28], and the estimation of signal parameters via rotational invariance [29] for the interpretation of higher resolutions were not considered to avoid loss of data input to the subsequent deep network.

B. COMPLEX-VALUED ATTENTION MODULE

In the case of a real-valued channel attention task [21] $f_{CA}^{\mathbb{R}}: \mathbf{A} \in \mathbb{R}^{H\times W\times C} \rightarrow \mathbf{X} \in \mathbb{R}^{H\times W\times C}$, the channel feature $\mathbf{v} \in \mathbb{R}^C$ can be obtained through a pooling operation for each channel $\mathbf{a}_c \in \mathbb{R}^{H\times W}$, where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots \mathbf{a}_C]$. To exploit the inter-dependencies between the channels from the representative channel feature, the scaling factor $\mathbf{s}_c$ can be determined by

$$\mathbf{s}_c = \delta\left(\mathrm{Conv}_U\left(\mathrm{ReLU}\left(\mathrm{Conv}_D\left(\mathbf{v}\right)\right)\right)\right) \quad (5)$$

where ReLU is a rectified linear unit activation [30]; δ is a sigmoid gating; and Conv_D and Conv_U are the convolution operations of 1×1 kernel implemented for channel down- and up-scaling, respectively, with a reduction ratio of r. Then, the scaling factor $\mathbf{s}_c$ is used to rescale the 3D real-valued tensor, $\mathbf{A}$.

$$\mathbf{X}_c = \mathbf{s}_c \cdot \mathbf{A}_c, \quad \forall c \in \{1,2,\cdots C\} \quad (6)$$

where $\mathbf{X}_c$ and $\mathbf{A}_c$ are the c-th channel feature maps of $\mathbf{X}$ and $\mathbf{A}$, respectively.

The overall flow of the proposed $\mathbb{C}$CA for a complex-valued 3D tensor, $\mathbf{D} = \mathbf{A} + i\mathbf{B} \in \mathbb{C}^{H\times W\times C}$, is shown in Fig. 3. For the $\mathbb{C}$CA task $f_{CA}^{\mathbb{C}}: \mathbf{D} \rightarrow \mathbf{Y} \in \mathbb{C}^{H\times W\times C}$ with a

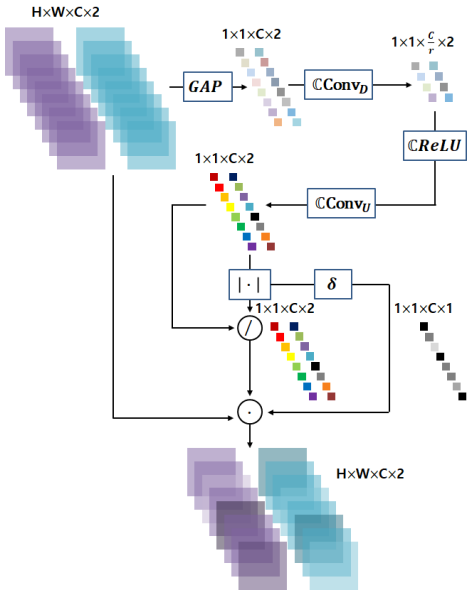

FIGURE 3. Overall structure of the proposed complex-valued channel attention module. For display purpose, each complex-valued tensor is represented by two real-valued tensors corresponding to real and imaginary parts.

feature map $\mathbf{d}_c \in \mathbb{C}^{H\times W}$, the channel feature $\mathbf{u} \in \mathbb{C}^C$ is first calculated using global average pooling as follows:

$$\mathbf{u} = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}\left\{\Re\left(\mathbf{d}_c(i,j)\right) + i\Im\left(\mathbf{d}_c(i,j)\right)\right\}. \quad (7)$$

To extract the inter-dependencies between channels, we applied down- and up-scaling convolution with a 1×1 convolution and a gating mechanism in-between similar to that of the real-valued task. The complex-valued generalization of convolution operation $\mathbb{C}$Conv can be defined [31] as

$$\mathbf{D} * \mathbf{w} = (\mathbf{A} * \mathbf{p} - \mathbf{B} * \mathbf{q}) + i\left(\mathbf{B} * \mathbf{p} + \mathbf{A} * \mathbf{q}\right). \quad (8)$$

where the convolution kernel is $\mathbf{w} = \mathbf{p} + i\mathbf{q}$. We exploit a complex-valued activation $\mathbb{C}$ReLU, similar to the gating mechanism, which is defined as

$$\mathbb{C}\mathrm{ReLU}(z) = \mathrm{ReLU}(\Re(z)) + i\mathrm{ReLU}(\Im(z)). \quad (9)$$

After the consecutive operation on the channel feature, the output $\hat{\mathbf{u}} \in \mathbb{C}^C$ can be written as

$$\hat{\mathbf{u}} = \mathbb{C}\mathrm{Conv}_U\left(\mathbb{C}\mathrm{ReLU}\left(\mathbb{C}\mathrm{Conv}_D\left(\mathbf{u}\right)\right)\right). \quad (10)$$

Inspired by the masking approaches [22]–[25], we optimized the polar-coordinate-wise cRM [25] that can produce a scaling factor $\mathbf{s}_c$ from $\hat{\mathbf{u}}$. This cRM imposes a non-linearity on the magnitude, while preserving the phase information as follows:

$$\mathbf{s}_c = \delta\left(\left|\hat{\mathbf{u}}_c\right|\right) \cdot e^{i\left(\hat{\mathbf{u}}_c/|\hat{\mathbf{u}}_c|\right)}. \quad (11)$$

The sigmoid gating $\delta(\cdot)$ bounds the magnitude of the scaling factor into a unit circle in the complex space. The corresponding phase information is obtained and passed through

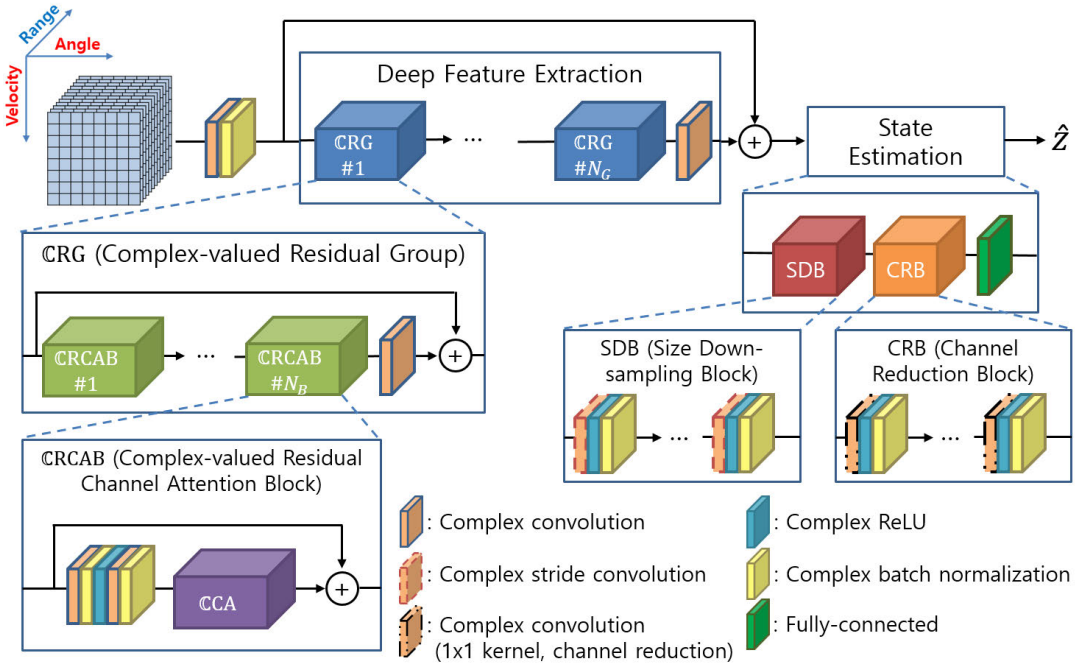

FIGURE 4. Overall network architecture for the ego-velocity estimation task.

the masking operation without loss. Similarly, the spatial attention module can be implemented through masking after extracting the representative feature $\hat{\mathbf{u}}_s \in \mathbb{C}^{H\times W}$ instead of $\hat{\mathbf{u}}$.

C. NETWORK ARCHITECTURE

The overall network architecture with the proposed ℂCA module is shown in Fig. 4. Because the naive stacking of the attention module would result in performance degradation, we coupled the attention module and residual learning, similar to numerous previous studies [8], [21], [32]–[35]. Specifically, we employed the REVEN architecture proposed in our previous study [8], which uses the body structure of RCAN [21] as the backbone.

The solution of EVE is based on the angle-velocity profile of static objects; thus, we reshaped the input spatial domain, viewed by the local receptive field, in the form of the angle-velocity profile:

$$\mathbf{S}_{in} = H_{ra}\left(\mathbf{S}\right), \tag{12}$$

where $H_{ra}\left(\cdot\right)$ denotes the data rearrangement operation. The overall process of network input data preparation is shown in Fig. 5. The convolution kernel processes the entire range, i.e., the channel of each local angle-velocity patch, via reshaping. This entire channel, a range profile, of the radar data requires the filtering of areas without objects, i.e., target detection. By applying the proposed ℂCA module, channels that contain objects can be emphasized (Fig. 6). Furthermore, the proposed E2E architecture can compensate for phenomena, such as range migration or phase wrapping, which can occur in high-speed maneuvering ego-platform situations [36].

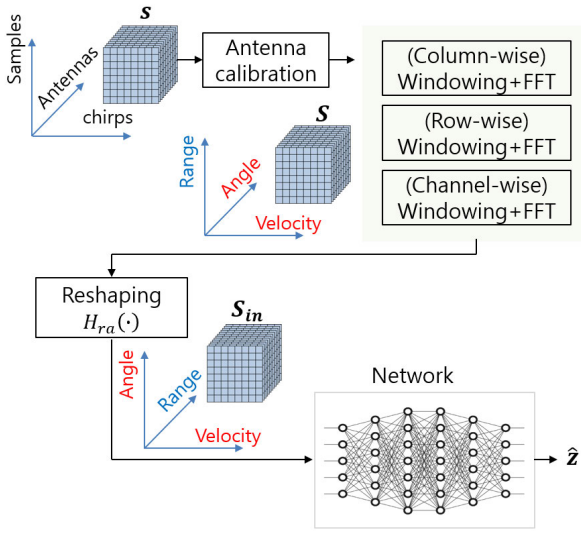

FIGURE 5. Data preparation scheme to use raw radar data as input for the neural network.

The network can be decomposed into three parts: the initial feature extraction (IFE), deep feature extraction (DFE), and state estimation. The deployment of one convolution layer is followed by batch normalization of the IFE part to extract the initial feature, S_{IF}, with the desired number of channels, as follows:

$$\begin{aligned}\mathbf{S}_{IF} &= H_{IFE}\left(\mathbf{S}_{in}\right)\\ &= \mathbb{C}\mathrm{BN}\left(\mathbb{C}\mathrm{Conv}\left(\mathbf{S}_{in}\right)\right)\end{aligned} \tag{13}$$

where $H_{IFE}\left(\cdot\right)$ is the IFE function, and ℂBN denotes the complex-valued batch normalization. Then, the obtained

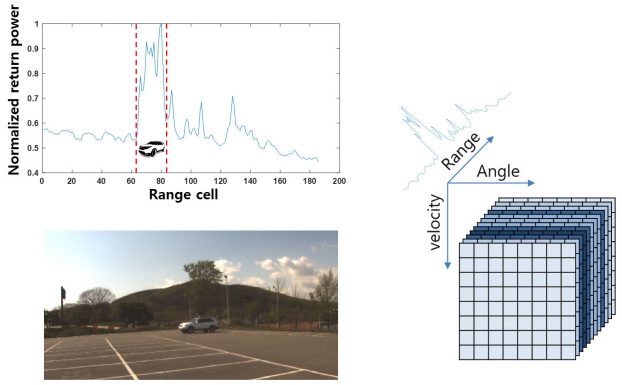

FIGURE 6. Left: Example of target detection in a range profile. Right: Illustration of the effect of the channel attention in the early stages of the network.

feature is fed into the DFE, where a residual in the residual structure is adopted. Skip connections in this residual make the deep network feasible [37]; furthermore, within the structure, the attention mechanism facilitates the extraction of the deep features required for state estimation.

Unlike the architecture of RCAN, which is a real-valued network for super-resolution (SR) tasks, all components of the network are used as complex-valued building blocks, including the attention module. Batch normalization (BN) [38] is also employed throughout the entire network to accelerate and stabilize training by reducing the internal covariance shift. This differs from other deep networks designed for image-to-image tasks, such as SR [21], [37] and image deblurring [39], because BN is often excluded from the network owing to the loss of scale information and removal of range flexibility.

The deep feature, S_{DF}, obtained through DFE is inputted with the initial feature, S_{IF}, through global skip connection and proceeds to the state estimation stage. This global shortcut facilitates the flow of information within the network by bypassing low-frequency information to the latter part of the network. The process of state estimation includes two main parts in terms of information compression and final estimation: the size down-sampling and channel reduction blocks, which use strided convolution and channel-reducing convolution with a 1×1 kernel and the last fully-connected layer. The ego-velocity state, $\hat{\mathbf{z}}$, is estimated from the two information sources, i.e., abundant low-frequency information and high-frequency details, as follows:

$$\hat{\mathbf{z}} = H_{SE}\left(S_{IF} + H_{DFE}\left(S_{IF}\right)\right) \tag{14}$$

where, $H_{DFE}(\cdot)$ and $H_{SE}(\cdot)$ denote the DFE and state estimation functions, respectively.

D. GEOMETRIC CONSTRAINTS

We added the states for the sensor into $\hat{\mathbf{z}}$ along with the ego-velocity states, i.e., velocity states for the ego-platform, and regularized network training using GCs that satisfy the relationship between them. The ego-velocity information we aim to estimate is the linear velocity, v_{long}, and angular velocity, ω, as shown in Fig. 7, which has a ground truth measured through a high-precision inertial measurement unit (IMU) with differential global positioning system (D-GPS) support. The sensor velocity, $\mathbf{v}_s = (v_x, v_y)$, mounting position, (l, b), and viewing direction, θ_s, are used as dummy variables without ground truths; they are used for regularization to assist training and reduce over-fitting.

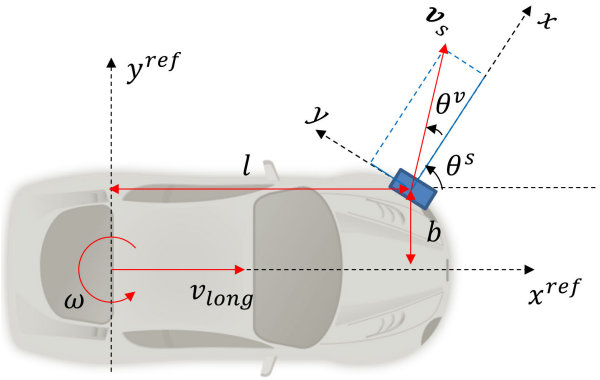

FIGURE 7. Illustration of the relationship between the reference and sensor coordinate systems.

As previously examined [9]–[11], [13], [40], the observed static target has a velocity that is directly opposite to the sensor velocity, $\mathbf{v}_s$. From the perspective of radars measuring the instantaneous radial velocity, $v_{r,i}$, of targets, the sensor velocity can be estimated using the bearing angle, θ_i, and the radial velocity relationships of two or more static targets, as follows:

$$\begin{bmatrix} v_{r,1} \\ \vdots \\ v_{r,N} \end{bmatrix} = \begin{bmatrix} \cos\theta_1 & \sin\theta_1 \\ \vdots & \vdots \\ \cos\theta_N & \sin\theta_N \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} \tag{15}$$

where N static targets are detected in one frame. The sensor velocity, $\mathbf{v}_s$, obtained from this relationship can be converted to the ego-velocity, if the sensor's mounting position and viewing direction are known, as follows:

$$\begin{aligned} v_{long} &= |\mathbf{v}_s| \cdot \left(\cos\left(\theta_s + \theta_v\right) + b \cdot \frac{\sin\left(\theta_s + \theta_v\right)}{l}\right) \\ \omega &= |\mathbf{v}_s| \cdot \frac{\sin\left(\theta_s + \theta_v\right)}{l} \end{aligned} \tag{16}$$

where θ_v is defined as

$$\theta_v = \tan^{-1}\left(\frac{v_y}{v_x}\right). \tag{17}$$

Herein, for simplicity, we assume no vehicle side-slip.

Because sensor velocity information estimated using the information of targets is transferred to the platform velocity information, we integrated dummy variables for GCs into the estimated state $\hat{\mathbf{z}}$ as follows:

$$\hat{\mathbf{z}} = \left[v_x, v_y, \theta_s, l, b, v_{long}, \omega\right]. \tag{18}$$

Then, we applied GCs, which impose the equality of (16), as a regularization loss, $\mathcal{L}_G$. The linear and angular components of the ego-velocities, v_{long} and ω, are optimized via the convex and continuous Huber loss [41], which is less sensitive to outliers compared with $\mathcal{L}_2$ loss. The final loss is defined by the sum of the Huber loss, $\mathcal{L}_H$, between the estimates and ground truth of the ego-velocity and regularization loss as follows:

$$\mathcal{L}_{total} = \mathcal{L}_H + \lambda \mathcal{L}_G. \tag{19}$$

By minimizing $\mathcal{L}_{total}$ with training samples in various situations, the network is implicitly forced to learn the mount information, l, b, and θ_s of the sensor because of regularization, which stabilizes the convergence of the training and improves performance.

IV. EXPERIMENTS AND DISCUSSION

To verify the performance of the proposed method on the EVE task performed with an automotive radar, we collected over 100,000 frames of radar data by driving through complex environments in a city. We verified that the E2E structure could be trained with complex-valued radar data to resolve the Doppler ambiguity problem without delicate correction efforts. To thoroughly evaluate the effectiveness of our network and its components, we also performed an ablation study and compared the performance of our model with that of previous approaches using our extensive database.

A. DATA ACQUISITION

The database comprised real-world measurements acquired by a radar mounted on the front bumpers of a vehicle. An automotive 79-GHz radar sensor with a valid sweep bandwidth of 512 MHz and an update-rate of 20 Hz was used. The range resolution was 29.3 cm, and the velocity resolution was 1.02 km/h using 64 chirps. In addition, the radar featured an array configuration with an aperture of 21.5λ, which corresponds to a 2.5° angle resolution.

For obtaining measurements at its ego-velocity, our vehicle was equipped with SPAN-CPT from NovaTel, which is a high-precision IMU with D-GPS support. The vehicle was driven at speeds of up to 120 km/h; the speed varied according to the surrounding urban environment. However, the maximum unambiguous velocity, v_{max}, of the radar was set to 32.59 km/h, at which the velocity ambiguity problem could easily occur in high-speed driving situations. Such a set-up is common in the time division multiplexing scheme, which is widely used to obtain the orthogonality between multiple transmitters. In the following section, we demonstrate that the proposed network can resolve the velocity ambiguity problem without hardware changes (e.g., overlapping elements [42]), changes in the transmission scheme (e.g., frequency shift [43], [44]), or additional signal processing steps [45].

Data were not collected more than twice in the same scene, and the database was divided into training, validation, and testing sets, of 83,253, 10,012, and 10,011 frames, respectively. The ego-velocity of each set was distributed over a wide range of velocities, as shown in Fig. 8. For most cases, the yaw rates are close to zero because of the urban driving environment, where cars are often driven straight forward. The testing set was distributed relatively evenly to include less static and fewer straight driving conditions.

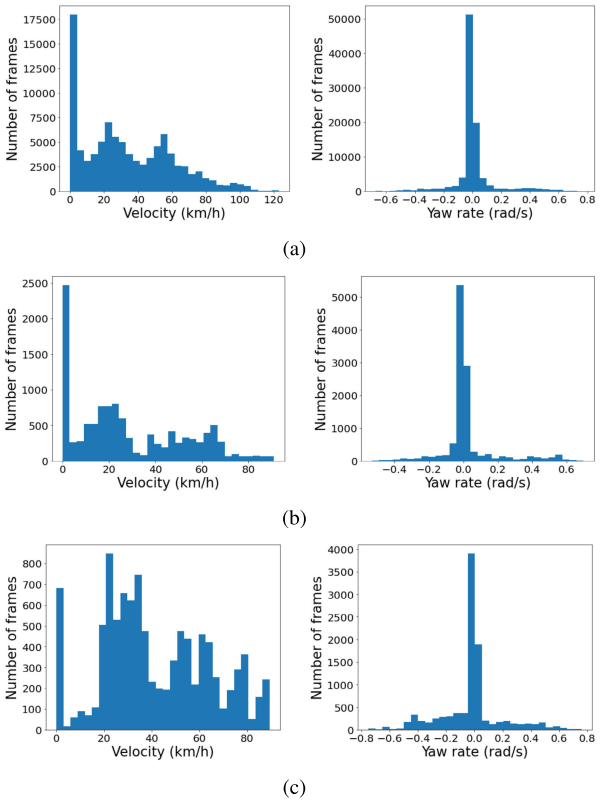

FIGURE 8. Distribution of velocity and yaw rate of ego-vehicle. (a) Training set, (b) Validation set, and (c) Testing set.

B. VELOCITY AMBIGUITY

As noted in Section III-D, the EVE task can be solved based on the instantaneous radial velocities of static targets. Although the detailed methods differ, the principle involves making estimations based on the angle-velocity curve (i.e., velocity profile) fitted by most targets, as shown in Fig. 9(a), and treating these targets as static targets.

However, the interval between chirp signals transmitted by the radar can set the maximum unambiguous velocity; in cases when targets move at a speed beyond this maximum, the phase due to the target's Doppler effect becomes wrapped, as shown in Figs. 9(b) and 9(c). For small excess of the speed, this phase-wrapping phenomenon only occurs around the antenna boresight. In such a situation, if most static targets are located at a high angle without wrapping, the estimation result may not be affected, as shown in Fig. 9(b). However, in severe cases, most measured values will be wrapped, seriously distorting the estimated curve, as shown in Fig. 9(c). Therefore, proper estimation is possible only when the unambiguous radial velocities are restored or when a delicate treatment of the wrapping phenomenon is performed as shown in Figs. 9(d) and 9(e).

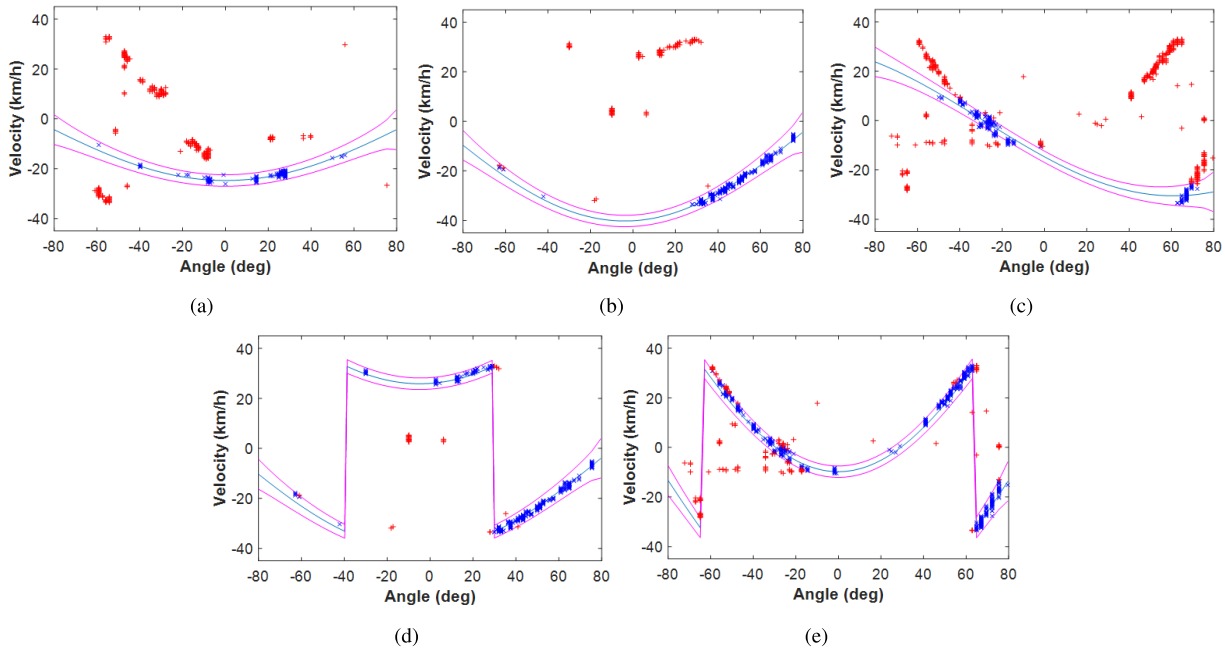

FIGURE 9. Examples of the angle-velocity plot. (a) No ambiguity, (b) Doppler ambiguity arising in boresight, (c) Severe ambiguity, and (d), (e) Ambiguity-resolved counterparts of (b) and (c). (Blue line: angle-velocity curve fitted by most targets; magenta line: supporting boundary; blue markers: targets fitting the blue line; red markers: targets outside the supporting boundary).

In practical situations, dealing with the Doppler ambiguity problem using conventional approaches [42]–[44] requires modifications in the hardware of modulation scheme accompanied by redundancy. In subsequent experimental comparisons, previous approach was evaluated only for situations without Doppler ambiguity. Although it is difficult to directly compare feature engineering-based techniques with this study, which has an E2E structure, seminal work [9] was used for performance comparison, and the parameters of the preceding target detection for the previous approach were optimized via training set.

Fig. 10 shows the results of the EVE task based on the testing set; the results are sorted and plotted in ascending order of the velocity. When the testing set is sorted in ascending order, the ego-velocity exceeds v_{max}, which is 32.59 km/h, at the 4138th frame. We first performed EVE using the method described in [9] (see Fig. 10(a)). In the early stage of the velocity ambiguity, the wrapping phenomenon starts to appear around the antenna boresight; however, data corresponding to the rest of the angle do not show this phenomenon, as shown in Fig. 9(b). If the remaining parts occupy a major portion, then the ego-velocity can still be accurately estimated, and the wrapping phenomenon, as shown in Fig. 9(d), can be reversely corrected. Through this experiment, we confirmed that accurate estimation can be achieved in unambiguous areas. Moreover, we verified that estimation performance does not deteriorate even if some velocity ambiguity exists. However, in cases with severe ambiguity, e.g., for velocity exceeding 50 km/h, the performance rapidly deteriorates.

Figure 10(a) shows that, even if the ego-velocity is within an unambiguous area, there are some frames with large estimation errors. Such errors are inevitable and occasionally appear in the previous technique because of the preceding results of target detection in situations where most detected targets are dynamic or with very few detected targets. However, our E2E approach enables the network to learn to handle the Doppler ambiguity problem; the proposed network can estimate the ego-velocity accurately, even when the driving velocity exceeds v_{max}, as shown in Fig. 10(b). Fig. 10(c) shows an enlarged section of the proposed results, which demonstrates the high precision achieved by the proposed method.

For a quantitative comparison, the mean absolute error (MAE) between the estimate and ground truth was calculated. Because no delicate techniques were added to handle the Doppler ambiguity problem, only estimates of 4,137 frames corresponding to unambiguous low-speed scenarios were used to calculate the MAE for the conventional method. The MAE calculation for the proposed method was computed using the entire test case of 10,011 frames; the results showed an accuracy of 0.3650 km/h against an accuracy of 0.5699 km/h for the conventional techniques in the low-speed case.

C. ABLATION STUDIES

To gain insight regarding the feasibility of the proposed method, we analyzed the effect of removing or adding components from the network. Specifically, we measured the effects of

- configuring the entire network with complex-valued building blocks,
- using the $\mathbb{C}$CA module that emphasizes the important channel without distorting the phase, and

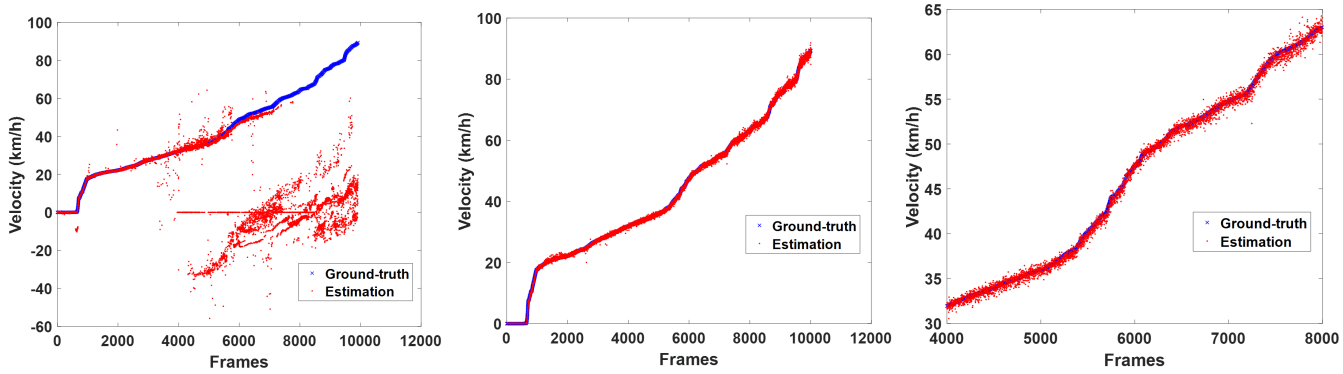

FIGURE 10. Velocity estimation results for the entire testing set (sorted in ascending order for visualization purposes). (a) Methods described in [9], (b) Proposed method, and (c) Enlarged view of (b).

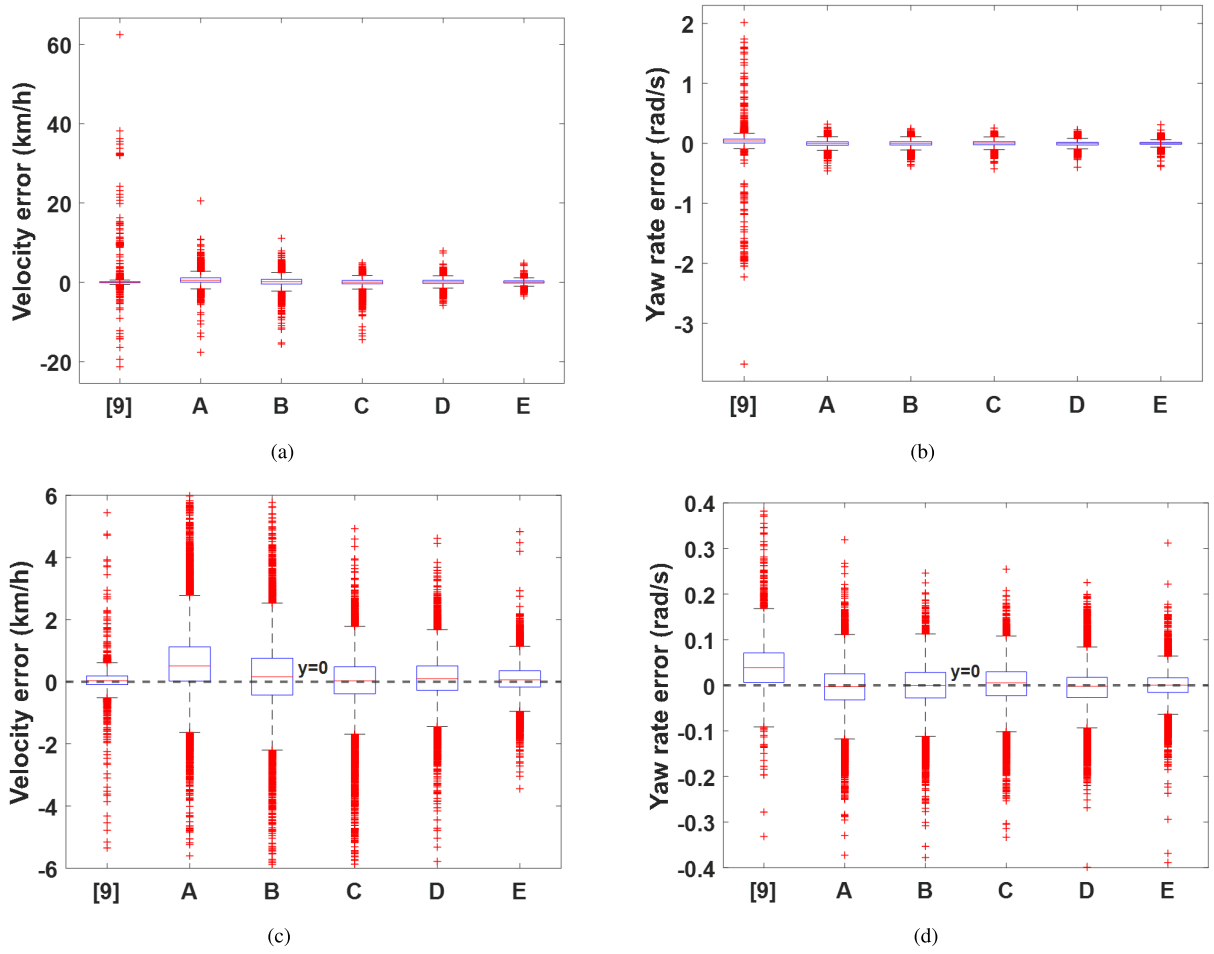

FIGURE 11. Box plot of the ablation study results using the proposed method and the testing set. (a) Velocity, (b) Yaw rate, and (c), (d) Enlarged counterparts of (a) and (b).

- applying regularization through constraints using geometric relationships between the sensor and ego-vehicle.

Our network was trained in the following environment. We adopted a mini-batch size of 128 with the maximum number of epochs set to 200. Our model was implemented with PyTorch and trained with Adam optimizer [46]. The learning rate was set to 0.001. Our network consisted of three $\mathbb{C}$RGs and four $\mathbb{C}$RCABs with the $\mathbb{C}$CA module in each block.

When the number of channels is the same, the $\mathbb{C}$Conv uses both real and complex kernels so it has twice the network capacity compared to its real-valued counterpart. In the case of RVNN, the real and imaginary parts of input are concatenated along the channel axis and the number of convolution kernels is doubled to match the level of capacity. In addition, the number of $\mathbb{C}$RCABs is doubled to compensate for the reduced capacity in the case of an

TABLE 1. MACs and Accuracy Results of the Ablation Study (Testing Set).

Model	CCA	CVNN	GC	MACs	MAE	
					v (km/h)	ω (rad/s)
[9]	·	·	·	·	0.5699	0.0846
A	✗	✗	✗	16.1G	0.9803	0.0408
B	✗	✓	✗	16.1G	0.8788	0.0389
C	✓	✗	✗	11.2G	0.6844	0.0358
D	✓	✓	✗	11.2G	0.5515	0.0331
E	✓	✓	✓	11.3G	0.3650	0.0235

attention-less network. Consequently, models A and B use higher multiply–accumulate operations (MACs) instead of eliminating an attention module from the network, as listed in Table 1.

Table 1 also lists the experimental results of the ablation study for different settings. Note that, in [9], estimations were only performed for unambiguous cases, and the accuracy of the velocity estimation is similar to that of case D. However, the error in the yaw rate estimation was much higher than that of the baseline case A. The proposed structure accurately estimated the yaw rate regardless of whether velocity ambiguity occurred. Comparing the test models for cases A-D, a significant improvement was observed when both $\mathbb{C}$CA and CVNN were applied, and $\mathbb{C}$CA exhibited superior performance. In model E, GC was applied, resulting in further significant performance improvement, i.e., the application of regularization based on geometrical relation had a considerable effect on minimizing the loss functions in a very high dimensional space.

From the box plots in Fig. 11, it is possible to determine the distribution of estimation errors and tendency of change in precision according to the selected model. The estimation performance of the method in [9] heavily relies on the preceding target detection step, and occasional frames with severe estimation errors were encountered, leading to a wide range of outliers in the box plots. In terms of the velocity error, the distribution range of outliers is significantly reduced by the proposed model, especially when CVNN and $\mathbb{C}$CA are used together, and the application of GC further improves both the accuracy and precision. Although the level of performance gain is slightly lower than that of the velocity, the yaw rate case achieved the highest performance for model E.

V. CONCLUSION

In this study, we propose a novel attention method for dealing with a complex-valued tensor whose amplitude and phase are related to the EM scattering of the target being observed. Our $\mathbb{C}A$ module was coupled with the REVEN architecture as a channel attention network; REVEN was designed as a CVNN, which is known to be particularly powerful for handling wave phenomena, such as EM waves and sound waves [47]. We also analyzed the geometric relation and leveraged this relation to regularize the training of the overall network. Our E2E structure relieved the burden of preprocessing dependencies of the EVE task, and the ego-velocity in the reference frame was obtained from the sensor velocity without sensor mounting information.

The experimental results demonstrated that the proposed method can accurately estimate the ego-velocity regardless of the occurrence of Doppler ambiguity. The ablation study also showed performance improvements both in terms of accuracy and precision, which resulted from the implementation of $\mathbb{C}$CA, CVNN, and GC. As a future endeavor, we aim to apply and verify our $\mathbb{C}$CA network to radar-based simultaneous localization and mapping.

REFERENCES

[1] 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.
[2] M. Steiner, O. Hammouda, and C. Waldschmidt, "Ego-motion estimation using distributed single-channel radar sensors," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Apr. 2018, pp. 1–4.
[3] B. Major, D. Fontijne, A. Ansari, R. T. Sukhavasi, R. Gowaikar, M. Hamilton, S. Lee, S. Grzechnik, and S. Subramanian, "Vehicle detection with automotive radar using deep learning on range-azimuth-Doppler tensors," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW)*, Oct. 2019.
[4] J. F. Tilly, S. Haag, O. Schumann, F. Weishaupt, B. Duraisamy, J. Dickmann, and M. Fritzsche, "Detection and tracking on automotive radar data with deep learning," in *Proc. IEEE 23rd Int. Conf. Inf. Fusion (FUSION)*, Jul. 2020, pp. 1–7.
[5] J. Lombacher, M. Hahn, J. Dickmann, and C. Wohler, "Potential of radar for static object classification using deep learning methods," in *IEEE MTT-S Int. Microw. Symp. Dig.*, May 2016, pp. 1–4.
[6] 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)*, Apr. 2019, pp. 1–6.
[7] M. S. Seyfioglu, A. M. Ozbayoglu, and S. Z. Gurbuz, "Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 4, pp. 1709–1723, Aug. 2018.
[8] H. W. Cho, S. Choi, Y.-R. Cho, and J. Kim, "Deep complex-valued network for ego-velocity estimation with millimeter-wave radar," in *Proc. IEEE Sensors*, Oct. 2020, pp. 1–4.
[9] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using Doppler radar," in *Proc. 16th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2013, pp. 869–874.
[10] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using multiple Doppler radars," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2014, pp. 1592–1597.
[11] M. Barjenbruch, D. Kellner, J. Klappstein, J. Dickmann, and K. Dietmayer, "Joint spatial- and Doppler-based ego-motion estimation for automotive radars," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2015, pp. 839–844.
[12] M. Rapp, M. Barjenbruch, K. Dietmayer, M. Hahn, and J. Dickmann, "A fast probabilistic ego-motion estimation framework for radar," in *Proc. Eur. Conf. Mobile Robots (ECMR)*, Sep. 2015, pp. 1–6.
[13] M. Rapp, M. Barjenbruch, M. Hahn, J. Dickmann, and K. Dietmayer, "Probabilistic ego-motion estimation using multiple automotive radar sensors," *Robot. Auto. Syst.*, vol. 89, pp. 136–146, Mar. 2017.
[14] C. D. Monaco and S. N. Brennan, "RADARODO: Ego-motion estimation from Doppler and spatial data in RADAR images," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 3, pp. 475–484, Sep. 2020.
[15] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[16] P. Biber and W. Strasser, "The normal distributions transform: A new approach to laser scan matching," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, vol. 3, 2003, pp. 2743–2748.
[17] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3104–3112.
[18] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: http://arxiv.org/abs/1406.1078

[19] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2019, pp. 7354–7363.
[20] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7794–7803.
[21] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 286–301.
[22] H. Erdogan, J. R. Hershey, S. Watanabe, and J. Le Roux, "Phase-sensitive and recognition-boosted speech separation using deep recurrent neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 708–712.
[23] D. S. Williamson, Y. Wang, and D. Wang, "Complex ratio masking for monaural speech separation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 24, no. 3, pp. 483–492, Mar. 2016.
[24] A. Ephrat, I. Mosseri, O. Lang, T. Dekel, K. Wilson, A. Hassidim, W. T. Freeman, and M. Rubinstein, "Looking to listen at the cocktail party," *ACM Trans. Graph.*, vol. 37, no. 4, p. 1–11, Aug. 2018, doi: 10.1145/3197517.3201357.
[25] H.-S. Choi, J.-H. Kim, J. Huh, A. Kim, J.-W. Ha, and K. Lee, "Phase-aware speech enhancement with deep complex U-Net," 2019, *arXiv:1903.03107*. [Online]. Available: http://arxiv.org/abs/1903.03107
[26] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.
[27] J. Bechter, F. Roos, and C. Waldschmidt, "Compensation of motion-induced phase errors in TDM MIMO radars," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 12, pp. 1164–1166, Dec. 2017.
[28] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
[29] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
[30] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. ICML*, 2010, pp. 807–814.
[31] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks," 2017, *arXiv:1705.09792*. [Online]. Available: http://arxiv.org/abs/1705.09792
[32] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, and H. Lu, "Dual attention network for scene segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3146–3154.
[33] Z. Shi, C. Chen, Z. Xiong, D. Liu, Z.-J. Zha, and F. Wu, "Deep residual attention network for spectral image super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops*, Sep. 2018, pp. 214–229.
[34] H. Ling, J. Wu, L. Wu, J. Huang, J. Chen, and P. Li, "Self residual attention network for deep face recognition," *IEEE Access*, vol. 7, pp. 55159–55168, 2019.
[35] A. Muqeet, M. T. B. Iqbal, and S.-H. Bae, "Hran: Hybrid residual attention network for single image super-resolution," *IEEE Access*, vol. 7, pp. 137020–137029, 2019.
[36] S. Neemat, O. Krasnov, F. van der Zwan, and A. Yarovoy, "Decoupling the Doppler ambiguity interval from the maximum operational range and range-resolution in FMCW radars," *IEEE Sensors J.*, vol. 20, no. 11, pp. 5992–6003, Jun. 2020.
[37] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 136–144.
[38] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: http://arxiv.org/abs/1502.03167
[39] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3883–3891.
[40] C. Grimm, T. Breddermann, R. Farhoud, T. Fei, E. Warsitz, and R. Haeb-Umbach, "Hypothesis test for the detection of moving targets in automotive radar," in *Proc. IEEE Int. Conf. Microw., Antennas, Commun. Electron. Syst. (COMCAS)*, Nov. 2017, pp. 1–6.
[41] P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in Statistics*. Cham, Switzerland: Springer, 1992, pp. 492–518.
[42] C. M. Schmid, R. Feger, C. Pfeffer, and A. Stelzer, "Motion compensation and efficient array design for TDMA FMCW MIMO radar systems," in *Proc. 6th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2012, pp. 1746–1750.
[43] K. Thurn, D. Shmakov, G. Li, S. Max, M.-M. Meinecke, and M. Vossiek, "Concept and implementation of a PLL-controlled interlaced chirp sequence radar for optimized range–Doppler measurements," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 10, pp. 3280–3289, Oct. 2016.
[44] W. Wang, J. Du, and J. Gao, "Multi-target detection method based on variable carrier frequency chirp sequence," *Sensors*, vol. 18, no. 10, p. 3386, Oct. 2018.
[45] F. Roos, J. Bechter, N. Appenrodt, J. Dickmann, and C. Waldschmidt, "Enhancement of Doppler unambiguity for chirp-sequence modulated TDM-MIMO radars," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Apr. 2018, pp. 1–4.
[46] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980
[47] A. Hirose, *Complex-Valued Neural Networks: Advances and Applications*, vol. 18. Hoboken, NJ, USA: Wiley, 2013.

HYUN-WOONG CHO received the B.S. and Ph.D. degrees in electrical engineering from the Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in February 2009 and August 2017, respectively. Since September 2017, he has been a Senior Researcher with the Machine Learning Lab, AI & SW Research Center, Samsung Advanced Institute of Technology (SAIT), Suwon-si, South Korea. His current research interests include automotive radar signal processing techniques, such as improved direction of arrival estimation, static target indication, ego-motion estimation, target recognition, and tracking.

SUNGDO CHOI received the B.S. degree in electrical engineering from Kyungbook National University, in 2004, and the Ph.D. degree in bio and brain engineering from KAIST, in 2011. He is currently a Principal Researcher with the Samsung Advanced Institute of Technology, South Korea. His research interests include radar signal processing and its application-based on machine learning.

YOUNG-RAE CHO received the B.S. degree in electrical engineering and the Ph.D. degree from the Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in 2012 and 2019, respectively. Since 2019, he has been working with the Samsung Advanced Institute of Technology (SAIT). His research interests include image processing, radar signal processing, adaptive signal processing, machine learning, and deep learning.

JONGSEOK KIM was born in Busan, South Korea. He received the M.Sc. degree in electronic engineering from Kyungpook National University, Daegu, South Korea, in 2000, and the Ph.D. degree in electronic engineering from Korea University, Seoul, South Korea, in 2010. He is currently a Principal Researcher with the Machine Learning Lab, Samsung Advanced Institute of Technology. His research interests include design and fabrication of IR sensor, acoustic sensor, mmWave active and passive sensors, and radar signal processing algorithms, having contributed, over 13 technical publications and 50 patents.