Data Augmentation for Synthetic Aperture Radar Target Recognition by Superimposing Real Target Signatures onto Statistically Simulated Clutter

Mykola Mazur[1]

1. Department of Software, Lviv Polytechnic National University, UKRAINE, Lviv, 12 S. Bandery Str., E-mail: mykola.mazur@lpnu.ua

Abstract ***To eliminate background bias in synthetic aperture radar target recognition without distorting radar shadow geometry, segmented target-and-shadow signatures are superimposed onto statistically simulated Weibull and K-distributed clutter via zero-memory nonlinear transformation. Evaluated in PyTorch on the MSTAR dataset across three convolutional networks, this method reaches 95.94% accuracy, matching real-clutter training.***

Keywords synthetic aperture radar, automatic target recognition, data augmentation, clutter simulation, target superimposition, Weibull distribution, K-distribution, MSTAR.

I. Introduction

Synthetic aperture radar (SAR) actively illuminates the ground with microwave pulses and records the amplitude of the reflected signal independently of daylight, cloud cover, smoke, or rain. Consequently, automatic vehicle recognition in SAR imagery is a key function of airborne reconnaissance systems. Because onboard computers of unmanned aerial vehicles operate under strict hardware and power limits [1], running heavy deep generative networks to filter background terrain reflections (known as radar clutter) during flight is impractical. Instead, robustness to changing terrain backgrounds must be built directly into a compact classification network during ground-based training.

The standard benchmark for SAR vehicle recognition is the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset, which contains cropped 128×128-pixel radar images (called target chips) of military vehicles recorded on uniform test-range soil. Because every vehicle class was captured on homogeneous ground, convolutional neural networks memorize the mean brightness of the surrounding terrain rather than the vehicle geometry-a problem known as background bias. Geng et al. [2] demonstrated that a ResNet-18 classifier achieving 97.2% accuracy on standard MSTAR images drops to 65.3% once background brightness is equalized across classes. In real operation, when a vehicle moves from open grass onto asphalt, forest edges, or urban blocks, such a classifier loses reliability.

In optical image processing, background overfitting is reduced either by geometric data augmentation (rotations and flips) or by empirical background replacement (often called clutter transfer in radar literature), where a segmented target is pasted onto a real background scene. For SAR imagery, both approaches face serious obstacles. First, rotating a radar frame by 90° or 180° breaks radar imaging physics: it rotates the dark radar shadow cast behind the vehicle relative to the fixed radar antenna line of sight, creating an impossible wave-scattering geometry. Second, empirical clutter transfer [3], [4] is limited by the scarcity of real radar background recordings (the public MSTAR archive provides only 50 clutter panoramas), while full three-dimensional electromagnetic simulation [5] requires proprietary 3D vehicle models.

To solve this problem, this paper proposes and experimentally validates a hybrid data augmentation method based on **superimposing real SAR target and shadow signatures onto statistically simulated clutter backgrounds**. Unlike empirical clutter transfer, which copies patches from a small set of real scenes, our method mathematically synthesizes an unlimited variety of spatially correlated clutter fields. **The main contributions of this work are:** (1) a four-block

augmentation pipeline that preserves both the bright metallic target reflection and its dark radar shadow while replacing the background with simulated clutter across six terrain categories; (2) a statistical clutter generator based on the zero-memory nonlinear transformation (ZMNL-a memoryless pixel-by-pixel mapping that reshapes the amplitude distribution without altering spatial spot sizes) of Weibull and compound K-distributions, verified by the Kolmogorov–Smirnov test; (3) an experimental proof across three neural networks (ResNet-18, EfficientNet-B0, and MobileNetV3-Small) that geometric rotations degrade SAR accuracy by 17.61–47.89 percentage points (pp) whereas the proposed superimposition raises ResNet-18 accuracy to 95.94%; and (4) a paired McNemar equivalence test proving that training on simulated clutter matches training on real radar clutter panoramas (96.85% vs. 97.10%, probability value $p = 0.4028$).

II. Related Work

Classical radar target detection relies on adaptive threshold algorithms known as Constant False Alarm Rate (CFAR) detectors, which estimate local clutter power inside a sliding window around a candidate pixel under the assumption of uniform background statistics [6], [7]. On heterogeneous landscapes such as urban areas, forest belts, or road boundaries, radar reflections exhibit sharp intensity bursts (heavy-tailed distributions), causing classical detectors to produce excessive false alarms. Deep generative denoising networks improve SAR classification [8], but their high computational cost conflicts with onboard processor constraints [1].

Offline background replacement shifts all computation to ground training. Existing clutter transfer methods cut the target out of a clean test-range image and blend it onto patches cropped from real clutter scenes [3], [4], remaining restricted by the 50 available MSTAR panoramas. Full electromagnetic domain randomization [5] removes this limit but requires closed 3D vehicle models and expensive ray-tracing simulators. In statistical radiophysics, terrain clutter amplitude is accurately described by the Weibull distribution for moderately heterogeneous surfaces and by the compound K-distribution for highly heterogeneous surfaces containing discrete scatterers [7], [9], [10]. Superimposing real target-and-shadow signatures onto mathematically generated, spatially correlated Weibull and K-distributed clutter fields provides a lightweight, physically valid alternative to both empirical background copying and 3D electromagnetic simulation.

III. Proposed Hybrid Augmentation Method

In a SAR image, each pixel with horizontal and vertical grid coordinates (x, y) stores the amplitude of the reflected microwave pulse in slant-range (distance from the radar antenna) and azimuth (flight-path) coordinates. The metallic hull of a vehicle forms a cluster of bright scattering centers, whereas the terrain directly behind the hull (away from the radar antenna) is shielded from microwave energy and forms a dark radar shadow. Rotating a radar image by 90° or 180° turns this shadow to the side or in front of the vehicle, disorienting the classifier. To diversify the terrain background without violating radar shadow geometry, the proposed method never rotates the target or shadow and replaces only the surrounding background. As shown in Fig. 1, the method operates in four sequential functional blocks.

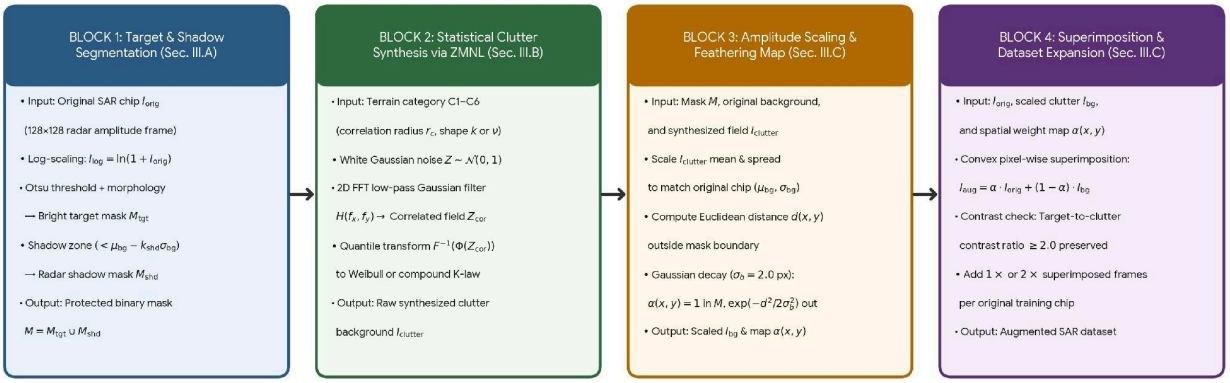

Fig. 1. Structural four-block diagram of the proposed SAR data augmentation method.

A. Block 1: Target and Radar Shadow Segmentation

Block 1 isolates both the bright metallic reflections of the vehicle and the dark radar shadow into a single protected binary mask M (where pixel value 1 denotes the protected vehicle-and-shadow region and 0 denotes the background) so that neither feature is overwritten by synthetic clutter. The four steps of this segmentation procedure are illustrated in Fig. 2. First, to compress the large dynamic range between bright metallic reflections and weak background terrain, the original 128×128 radar amplitude chip I_{orig} (subscript "orig" denotes original) is converted to a logarithmic scale, $I_{log} = \ln(1 + I_{orig})$ (Fig. 2a), where $\ln(\cdot)$ is the natural logarithm and adding 1 ensures that zero-amplitude pixels map to zero rather than negative infinity. Second, Otsu's automatic thresholding method [11] separates the bright target scatterers from the darker background on I_{log}, and morphological closing (3×3-pixel kernel) followed by opening removes isolated noise pixels to form the binary target mask M_{tgt} (subscript "tgt" denotes target, Fig. 2b). Third, using the background pixels outside M_{tgt}, the sample mean background brightness μ_{bg} and standard deviation σ_{bg} (subscript "bg" denotes background) are calculated. Dark pixels satisfying $I_{log} < \mu_{bg} - k_{shd} \cdot \sigma_{bg}$ adjacent to the vehicle hull are selected via connected-component filtering to form the radar shadow mask M_{shd} (subscript "shd" denotes shadow, Fig. 2c). Here $k_{shd} = 1.35$ is an empirically determined shadow threshold coefficient: requiring a pixel to be at least 1.35 standard deviations darker than the mean background μ_{bg} reliably separates the deep geometric radar shadow from ordinary dark background fluctuations. Fourth, taking the logical union ($\cup$) of both binary masks produces the combined protected mask $M = M_{tgt} \cup M_{shd}$ (Fig. 2d).

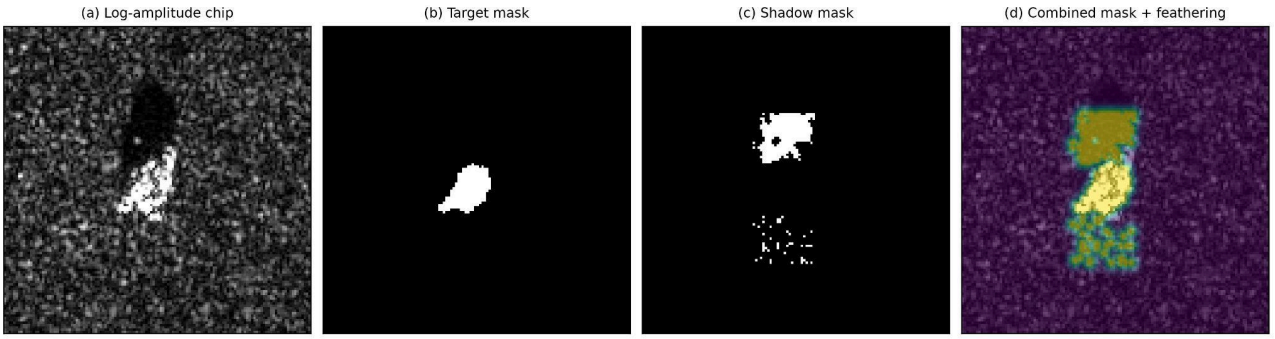

Fig. 2. Step-by-step visualization of Block 1 and Block 3 on a 128×128 SAR chip: (a) log-amplitude image I_{log}; (b) segmented target mask M_{tgt}; (c) segmented radar shadow mask M_{shd}; (d) combined protected mask M with smooth Gaussian feathering weights $\alpha(x, y)$.

B. Block 2: Statistical Synthesis of Correlated Clutter

Real terrain clutter has two physical properties: spatial correlation between neighboring pixels (caused by finite-sized objects such as tree crowns or buildings) and an asymmetric heavy-tailed amplitude distribution. In Block 2, both properties are reproduced mathematically using the zero-memory nonlinear transformation (ZMNL) algorithm in three sequential steps. First, a 128×128 matrix Z of independent standard Gaussian random noise with zero mean and unit variance, denoted as $N(0, 1)$, is generated. Second, Z is transformed into the two-dimensional spatial frequency domain (f_x, f_y) via the Fast Fourier Transform (FFT), multiplied by a low-pass Gaussian filter $H(f_x, f_y) = \exp(-2\pi^2 r_c^2(f_x^2 + f_y^2))$, transformed back via the inverse FFT, and normalized to zero mean and unit variance to obtain the spatially correlated Gaussian field Z_{cor} (subscript "cor" denotes correlated). Here r_c (subscript "c" denotes correlation) is the spatial correlation radius in pixels, which defines the characteristic radius of terrain clutter clusters (such as tree canopies or building blocks). Third, a memoryless pixel-by-pixel quantile transformation maps Z_{cor} into the desired radar amplitude distribution according to Eq. (1):

$$I_{clutter}(x, y) = F^{-1}(\Phi(Z_{cor}(x, y))), \quad (1)$$

where $I_{clutter}(x, y)$ is the synthesized raw clutter amplitude at pixel (x, y). The physical logic of Eq. (1) is straightforward: applying the standard normal cumulative distribution function $\Phi(\cdot)$ first maps the bell-shaped Gaussian values Z_{cor} onto the uniform probability interval (0, 1), after which the inverse cumulative distribution (quantile) function $F^{-1}(\cdot)$ of the chosen radar clutter law reshapes those probabilities into the required heavy-tailed terrain amplitude distribution. As listed in Table 1, six typical terrain clutter categories (denoted C1–C6, where "C" stands for Clutter category) are parameterized based on standard radiophysical literature [7], [9]. Categories C1–C3 use the compound K-distribution-defined by the homogeneity shape parameter v (smaller v produces sharper intensity bursts from discrete urban or forest scatterers) and the standard multi-look averaging count $L = 4$ (number of averaged radar sub-apertures)-while categories C4–C6 use the Weibull distribution defined by the shape parameter k (as k approaches 2.0 in C6, the clutter approaches homogeneous Rayleigh-like open grass).

Table 1

Parametrization of the Six Simulated Radar Clutter Categories C1–C6 (Block 2)

Clutter Code	Simulated Terrain Surface Type	Distribution Law	Shape Parameter (v or k)	Correlation Radius r_c (px)
C1	Dense urban / built-up area	Compound K-law	$v = 1.2$ (heavy tails)	2.5
C2	Forest belts and isolated tree clusters	Compound K-law	$v = 2.5$ (moderate tails)	4.0
C3	Roads and linear man-made structures	Compound K-law	$v = 4.0$ (mild tails)	2.0
C4	Natural terrain boundaries and shorelines	Weibull law	$k = 1.6$ (heterogeneous)	3.0

C5	Uniform artificial surfaces (asphalt)	Weibull law	$k = 1.9$ (near-uniform)	1.0
C6	Open plain ground and grass	Weibull law	$k = 2.0$ (homogeneous)	1.5

Fig. 3 compares illustrative realizations of simulated clutter categories C1, C2, and C6 (Fig. 3a–c) against a 128×128 real radar clutter patch cropped from a target-free MSTAR background panorama (archive file identifier HB06158, Fig. 3d). The probability density histograms in Fig. 3e show that simulated C6 open-ground clutter closely matches the empirical amplitude distribution of the real rural radar scene, while urban category C1 exhibits a heavier upper tail corresponding to bright building reflections. For a square frame of width $W = 128$ pixels, the 2D FFT computational complexity is $O(W^2 \log W)$, requiring only ~15 ms per frame on an Intel Core i7 processor.

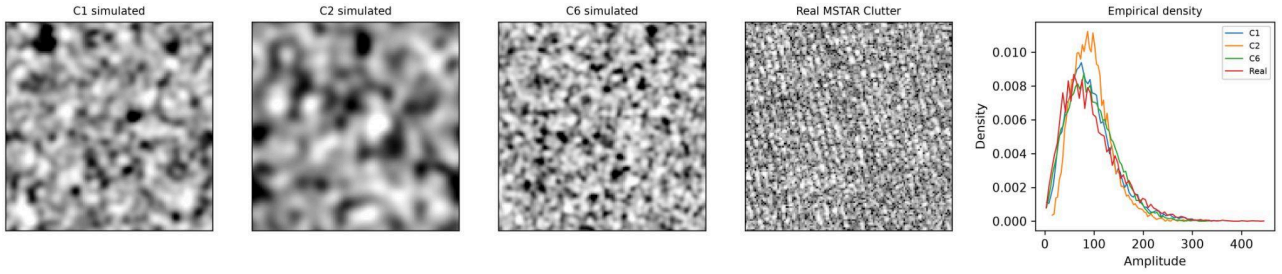

Fig. 3. Visual and statistical comparison of simulated clutter fields (a–c) versus a real MSTAR clutter patch from scene HB06158 (d) and their normalized amplitude probability density histograms (e).

C. Blocks 3 and 4: Amplitude Scaling, Feathering, and Target Superimposition

Directly pasting the segmented mask M onto the raw clutter field I_{clutter} would create an unnatural brightness jump and sharp pixel boundaries. In Block 3, the raw synthesized field I_{clutter} is first linearly scaled so that its mean and spread match the background mean μ_{bg} and standard deviation σ_{bg} of the original chip according to Eq. (2):

$$I_{\text{bg}}(x, y) = \mu_{\text{bg}} + ((I_{\text{clutter}}(x, y) - \mu_{\text{clutter}}) / \sigma_{\text{clutter}}) \cdot \sigma_{\text{bg}} , \quad (2)$$

where μ_{clutter} and σ_{clutter} are the sample mean and standard deviation of the raw synthesized field I_{clutter}, and I_{bg} is the resulting radiometrically scaled background. This scaling ensures that the target-to-clutter contrast ratio (the ratio of mean target amplitude inside M_{tgt} to mean background amplitude outside M) remains above 2.0 across all categories C1–C6, preserving the physical dominance of metallic reflections over terrain clutter. Next, to eliminate sharp step-boundaries around the mask contour, a continuous spatial feathering weight map $\alpha(x, y)$ ranging from 0 to 1 is computed from the Euclidean distance $d(x, y)$ (in pixels) between any background pixel (x, y) and the nearest boundary pixel of the protected mask M via Eq. (3):

$$\alpha(x, y) = 1 \text{ inside mask } M; \quad \alpha(x, y) = \exp(-d(x, y)^2 / (2\sigma_b^2)) \text{ outside mask } M, \quad (3)$$

where $\sigma_b = 2.0$ pixels (subscript "b" denotes boundary blending) is the Gaussian transition width, chosen so that the weight $\alpha(x, y)$ decays smoothly from 1 to 0 across a narrow 4–5-pixel rim outside the mask (visualized in Fig. 2d). Finally, in Block 4, the augmented radar image I_{aug} (subscript "aug" denotes augmented) is obtained by pixel-wise convex superimposition of the original chip I_{orig} and the scaled simulated background I_{bg} according to Eq. (4):

$$I_{aug}(x, y) = \alpha(x, y) \cdot I_{orig}(x, y) + (1 - \alpha(x, y)) \cdot I_{bg}(x, y). \quad (4)$$

IV. Experimental Setup

A. MSTAR Dataset Grouping and Superimposed Image Examples

Experiments were conducted on the 10-class MSTAR dataset, which comprises standard military and engineering vehicles (tracked main battle tanks, infantry fighting vehicles, self-propelled artillery, wheeled armored personnel carriers, bulldozers, and cargo trucks). In radar recognition benchmarks, training and testing sets are split by the radar depression angle (the vertical viewing angle of the radar beam relative to the horizontal ground plane). Frames recorded at a 17° depression angle form the development set (3,671 chips, divided into 3,303 training and 368 validation frames), whereas frames recorded at a 15° depression angle form the independent test set (3,203 frames). Table 2 organizes all 6,874 verified vehicle chips into physical vehicle groups. In addition, 50 full-size real MSTAR clutter panoramas (1784×1478 pixels, containing no vehicles) were used as an empirical reference baseline.

Table 2

Grouping and Sample Distribution of the 10-Class MSTAR Dataset by Depression Angle

Vehicle Group	Class Code	Vehicle Type Description	Train (17°)	Val (17°)	Test (15°)	Total
Main Battle Tanks	T-62	Medium main battle tank	269	30	273	572
Main Battle Tanks	T-72	Heavy main battle tank	209	23	196	428
Infantry Fighting Vehicles	BMP-2	Tracked infantry fighting vehicle	210	23	195	428
Artillery & Air Defense	2S1	122-mm self-propelled howitzer	269	30	274	573
Artillery & Air Defense	ZSU-23-4	Self-propelled anti-aircraft gun	269	30	274	573
Wheeled Armored Vehicles	BRDM-2	Armored reconnaissance scout car	268	30	274	572
Wheeled Armored Vehicles	BTR-60	8×8 wheeled armored personnel carrier	230	26	195	451
Wheeled Armored Vehicles	BTR-70	8×8 wheeled armored personnel carrier	210	23	196	429

Engineering & Logistics	D7	Heavy tracked military bulldozer	269	30	274	573
Engineering & Logistics	ZIL-131	6×6 general-purpose cargo truck	1,100	123	1,052	2,275
Total Benchmark	**10 Classes**	**All Verified Target Chips**	**3,303**	**368**	**3,203**	**6,874**

Fig. 4 shows examples of target chips generated by the proposed superimposition pipeline (Fig. 4c–e) compared against the original MSTAR chip (Fig. 4a) and a classical 90° rotation (Fig. 4b). While the 90° rotation shifts the dark radar shadow upward-right and violates the downward radar illumination geometry, the superimposed chips preserve the true bottom-right shadow orientation across simulated urban (C1), forest (C2), and grass (C6) backgrounds.

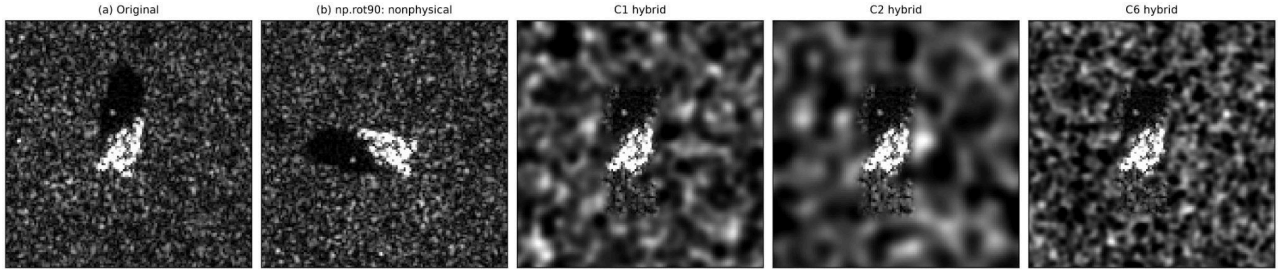

Fig. 4. Comparison of augmentation results on a 128×128 MSTAR chip: (a) original chip; (b) invalid 90° rotation with shifted shadow; (c–e) proposed superimposition onto simulated C1, C2, and C6 clutter.

B. Neural Network Architectures and Training Regimes

Three convolutional neural networks covering different computational budgets were trained in PyTorch: ResNet-18 (11.7 million parameters [12]), EfficientNet-B0 (4.0 million parameters [13]), and lightweight MobileNetV3-Small (1.5 million parameters [14]). Each network was trained for 15 epochs (Adam optimizer, batch size 64, NVIDIA GeForce RTX 4050 GPU) under four training dataset regimes (where N denotes the total number of training images): (1) **Baseline** (N = 3,303 original 17° chips); (2) **Classical Geometric Rotation** (N = 3,303 chips augmented with random 90°/180°/270° rotations, flips, and Gaussian noise); (3) **Proposed Superimposition (1×)** (N = 6,606 chips: each original chip paired with one superimposed frame randomly drawn from C1–C6); and (4) **Proposed Superimposition (2×)** (N = 9,909 chips: each original chip paired with two superimposed frames of different clutter categories). Model performance on the independent 15° test set (N_{test} = 3,203 chips) was evaluated using overall classification Accuracy (%) and the macro-averaged F1-score (Macro-F1, %, defined as the unweighted average of per-class harmonic means of precision and recall to account for class imbalance).

V. Results and Discussion

A. Statistical Verification and Augmentation Strategy Comparison

First, the statistical accuracy of the ZMNL clutter generator (Block 2) was verified on decorrelated 256×256 fields using the Kolmogorov–Smirnov (KS) goodness-of-fit test against theoretical Weibull and compound K-distributions. Across all six categories C1–C6, the maximum vertical distance D_{KS} between the empirical and theoretical cumulative distribution functions satisfied $D_{KS} \leq 0.0125$ with probability values p ranging from 0.4983 to 0.9966 (all well above the 5% significance threshold $p = 0.05$), confirming that the synthesized clutter accurately follows the theoretical distributions. Second, Table 3 and Fig. 5a summarize the classification performance on the 15° test set (N_{test} = 3,203 chips) across the four training regimes.

Table 3

Test Classification Performance at 15° Depression Angle (N_{test} = 3,203 Chips) Across Training Regimes

Architecture	Parameters / Size	Training Dataset Regime	Train *N*	Accuracy (%)	Macro-F1 (%)	Change (pp)
ResNet-18	11.7M (44.7 MB)	Baseline (Original chips)	3,303	94.63	94.57	-
ResNet-18	11.7M (44.7 MB)	Classical Geometric Rotation	3,303	77.02	76.65	-17.61 pp
ResNet-18	11.7M (44.7 MB)	Proposed Superimposition (1×)	6,606	95.47	95.21	+0.84 pp
ResNet-18	11.7M (44.7 MB)	**Proposed Superimposition (2×)**	**9,909**	**95.94**	**95.56**	**+1.31 pp**
EfficientNet-B0	4.0M (15.6 MB)	**Baseline (Original chips)**	**3,303**	**95.69**	**95.32**	**-**
EfficientNet-B0	4.0M (15.6 MB)	Classical Geometric Rotation	3,303	76.12	74.48	-19.57 pp
EfficientNet-B0	4.0M (15.6 MB)	Proposed Superimposition (1×)	6,606	93.13	91.93	-2.56 pp
EfficientNet-B0	4.0M (15.6 MB)	Proposed Superimposition (2×)	9,909	91.70	90.27	-3.99 pp
MobileNetV3-Small	1.5M (5.9 MB)	Baseline (Original chips)	3,303	89.35	87.81	-

MobileNetV3-Small	1.5M (5.9 MB)	Classical Geometric Rotation	3,303	41.46	34.46	-47.89 pp
MobileNetV3-Small	1.5M (5.9 MB)	**Proposed Superimposition (1×)**	**6,606**	**89.76**	**88.59**	**+0.41 pp**
MobileNetV3-Small	1.5M (5.9 MB)	Proposed Superimposition (2×)	9,909	85.11	82.77	-4.24 pp

As shown in Table 3, classical geometric rotation causes a severe accuracy drop across all three models: -17.61 percentage points (pp) on ResNet-18, -19.57 pp on EfficientNet-B0, and -47.89 pp on MobileNetV3-Small, confirming that rotating SAR frames destroys physical shadow orientation. Conversely, the proposed clutter superimposition increases ResNet-18 accuracy monotonically to 95.47% (1×) and 95.94% (2×), and improves MobileNetV3-Small to 89.76% at 1× expansion. Because the 15° test set was recorded on the same homogeneous test range as the 17° training set, the unaugmented Baseline benefits artificially from memorizing the uniform test-range soil; consequently, compact models perform best with moderate 1× clutter expansion, whereas deeper networks such as ResNet-18 benefit fully from 2× expansion.

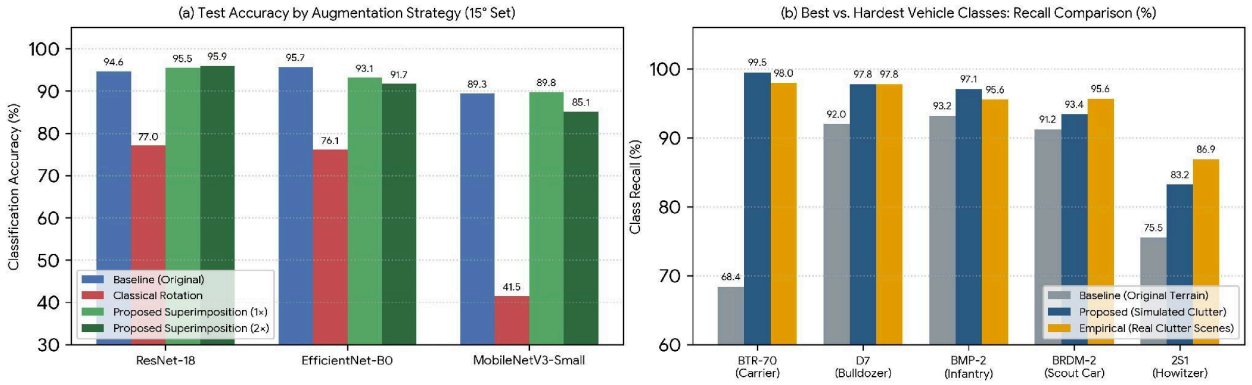

Fig. 5. Experimental results: (a) test accuracy across the four training regimes; (b) per-class recall for the best-recognized and hardest vehicle classes.

B. Analysis of Best- and Worst-Recognized Vehicle Classes and Real Clutter Equivalence

Rather than listing all ten classes individually, Figs. 5b and 6 compare the best-recognized and hardest-to-recognize vehicle categories after clutter superimposition. The highest per-class recall (percentage of true class samples correctly identified) is achieved on the BTR-70 wheeled armored carrier (99.49%, a gain of +31.12 pp over the 68.37% baseline on MobileNetV3-Small), the D7 heavy bulldozer (97.81%, +5.84 pp), and the BMP-2 infantry fighting vehicle (97.10%, +3.91 pp). As shown in Fig. 6 (panels a and b), the massive front bulldozer blade of the D7 and the elongated hull-to-shadow profile of the BTR-70 provide distinct geometric features once background bias is removed. The hardest classes to recognize are the 2S1 self-propelled howitzer (83.21% recall vs. 75.55% baseline) and the BRDM-2 compact scout car (93.43% vs. 91.24% baseline). As illustrated in Fig. 6 (panels c and d), the small hull of the BRDM-2 produces weaker reflections close to surrounding clutter spikes, while the tracked chassis and turret silhouette of the 2S1 howitzer closely resemble the T-72 main battle tank at oblique viewing angles.

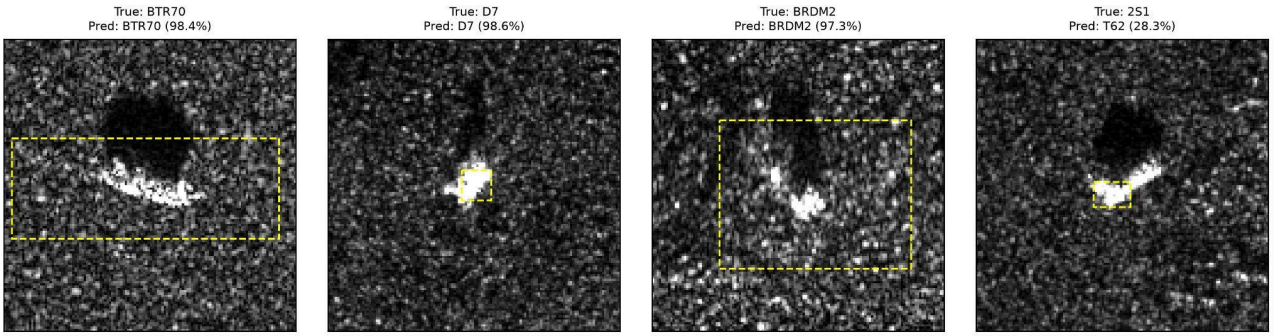

Fig. 6. Representative 15° test SAR chips for the best-recognized classes (BTR-70, D7) and hardest classes (BRDM-2, 2S1) showing target bounding boxes, true labels, predicted labels, and physical scattering features.

Finally, Table 4 compares two ResNet-18 classifiers trained on identical dataset sizes (N = 9,909): one augmented with the proposed ZMNL simulated clutter (C1–C6) and one augmented with patches cropped from the 50 real MSTAR clutter panoramas, evaluated on the N_{test} = 3,203 test chips using McNemar's paired test [15]. In this paired comparison, both models classify 3,071 chips correctly and 62 chips incorrectly, while n_{10} = 31 chips are recognized correctly only by the simulation-trained model and n_{01} = 39 chips only by the real-clutter model. From these discordant counts (n_{10} and n_{01}), McNemar's chi-square statistic with continuity correction is $\chi^2 = (|n_{01} - n_{10}| - 1)^2 / (n_{01} + n_{10}) = 0.70$, which corresponds to $p = 0.4028$.

Table 4

Paired Comparison of Training on Simulated Clutter vs. Real MSTAR Clutter Panoramas (N_{test} = 3,203)

Evaluation Indicator	Proposed Simulated Clutter	Real MSTAR Clutter	Statistical Equivalence Verdict
Overall Test Accuracy (%)	96.85% (3,102 / 3,203)	97.10% (3,110 / 3,203)	Difference: -0.25 pp
Discordant Correct Predictions	n_{10} = 31 chips (sim only)	n_{01} = 39 chips (real only)	3,071 both correct / 62 both wrong
95% Confidence Interval	-	-	[-0.76 pp; +0.26 pp] (within ±1%)
McNemar Paired Test (χ^2, p)	$\chi^2 = 0.70$	$p = 0.4028$	Not significant ($p \gg 0.05$)
Onboard Clutter Filtering Delay	0.0 ms (offline synthesis)	0.0 ms (offline cropping)	GPU inference: 3.73 ms (268 fps)

As Table 4 confirms, the classifier trained on simulated clutter reaches 96.85% accuracy against 97.10% for real clutter panoramas. Because p = 0.4028 is far above 0.05 and the 95% confidence interval of the accuracy difference ([-0.76 pp; +0.26 pp]) lies entirely within ±1.0 pp, mathematical clutter synthesis is statistically equivalent to real clutter scenes while adding zero onboard filtering latency (3.73 ms/frame or 268 frames/s on GPU; 8.14 ms/frame or 123 frames/s on CPU for MobileNetV3-Small).

VI. Conclusion

A physical-geometry-preserving data augmentation pipeline for SAR target recognition was developed by superimposing segmented target and radar shadow signatures onto statistically simulated Weibull and K-distributed clutter across six terrain categories. The ZMNL clutter generator is statistically adequate ($D_{KS} \leq 0.0125$, $p \geq 0.4983$). Whereas classical geometric rotations reduce recognition accuracy by 17.61–47.89 pp by breaking radar shadow orientation, the proposed superimposition preserves physical shadow geometry, increases ResNet-18 accuracy to 95.94%, and achieves statistical equivalence to training on real radar clutter panoramas (96.85% vs. 97.10%, McNemar $p = 0.4028$) without adding onboard inference latency.

References

[1] McEnroe, P., Wang, S., & Liyanage, M. (2022). A survey on the convergence of edge computing and AI for UAVs: Opportunities and challenges. *IEEE Internet of Things Journal*, *9*(17), 15403–15428. DOI: 10.1109/JIOT.2022.3176400

[2] Geng, Z., Xu, Y., Wang, B.-N., Yu, X., Zhu, D.-Y., & Zhang, G. (2023). Target recognition in SAR images by deep learning with training data augmentation. *Sensors*, *23*(2), 941. DOI: 10.3390/s23020941

[3] Zhang, M., et al. (2023). Mix MSTAR: A synthetic benchmark dataset for multi-class rotation vehicle detection in large-scale SAR images. *Remote Sensing*, *15*(18), 4558. DOI: 10.3390/rs15184558

[4] Denton, A. W. (2024). *Data augmentation for synthetic aperture radar using alpha blending and deep layer training* (Technical Report). Naval Postgraduate School. URL: https://calhoun.nps.edu/

[5] Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P. (2017). Domain randomization for transferring deep neural networks from simulation to the real world. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 23–30). IEEE. DOI: 10.1109/IROS.2017.8202133

[6] Ai, J., Cao, Z., & Xing, M. (2020). An adaptive-trimming-depth based CFAR detector of heterogeneous environment in SAR imagery. *Remote Sensing Letters*, *11*(8), 730–738. DOI: 10.1080/2150704X.2020.1763501

[7] Gao, S., & Liu, H. (2022). Performance comparison of statistical models for characterizing sea clutter and ship CFAR detection in SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 7414–7430. DOI: 10.1109/JSTARS.2022.3202965

[8] Remusati, H., Le Caillec, J.-M., Schneider, J.-Y., Petit-Frere, J., & Merlet, T. (2024). Generative adversarial networks for SAR automatic target recognition and classification models enhanced explainability. *Remote Sensing*, *16*(14), 2569. DOI: 10.3390/rs16142569

[9] Jao, J. K. (1984). Amplitude distribution of composite terrain radar clutter and the K-distribution. *IEEE Transactions on Antennas and Propagation*, *32*(10), 1049–1062. DOI: 10.1109/TAP.1984.1143200

[10] Hamidi, S. (2024). *Target detection in sea clutter with application to spaceborne SAR imaging*. arXiv Preprint, arXiv:2409.02155. URL: https://arxiv.org/abs/2409.02155

[11] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. DOI: 10.1109/TSMC.1979.4310076

[12] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). IEEE. DOI: 10.1109/CVPR.2016.90

[13] Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning (ICML)* (pp. 6105–6114). PMLR.

[14] Howard, A., et al. (2019). Searching for MobileNetV3. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 1314–1324). IEEE. DOI: 10.1109/ICCV.2019.00140

[15] McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, *12*(2), 153–157. DOI: 10.1007/BF02295996