UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

VISION TRANSFORMER MODEL FOR CLASSIFICATION OF AERIAL TARGETS USING NON-COHERENT X-BAND RADAR

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

AHMED AHMED
Norman, Oklahoma
2026

VISION TRANSFORMER MODEL FOR CLASSIFICATION OF AERIAL TARGETS USING NON-COHERENT X-BAND RADAR

A THESIS APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Justin Metcalf, Chair

Dr. Russell Kenney

Dr. Choon Yik Tang

This thesis is dedicated to the memory of my grandmother Fatheya

Acknowledgments

Foremost, I express my profound gratitude to Dr. Justin Metcalf for his unwavering support and the countless opportunities he has provided during my years at the Advanced Radar Research Center. His guidance has shaped my technical growth in radar systems, signal processing, and machine learning, while his example has shaped my character. Dr. Metcalf has shown me not only how to conduct rigorous research but also how to lead with integrity, curiosity, and kindness, making every step of this journey both challenging and rewarding.

I owe a lifelong debt of thanks to my mother, Samiha AbuElhamed. Her sacrifices and unfaltering belief in me gave a fifteen-year-old the courage to leave Egypt, pursue an education in Hong Kong, and ultimately embark on a Bachelor, and Master's at the University of Oklahoma. To my father, Abdallah Mohamed, thank you for demonstrating through action rather than words what it means to be honest, trustworthy, and relentlessly hard-working. Their love and example have been the steady foundation on which all my achievements rest.

This thesis is as much theirs as it is mine.

Table of Contents

List of Tables

List of Figures

Abstract

Small unmanned aerial vehicles can exhibit radar cross sections and bulk radial velocities comparable to those of birds, limiting discrimination based on range, received amplitude, and translational Doppler alone. Micro-Doppler signatures provide complementary information because wing flapping and rotor motion produce target-dependent modulation in the time–frequency domain. This thesis presents three related contributions. First, a leakage-referenced coherent-on-receive (COR) processing framework is developed for intermediate-frequency data collected with a Furuno FAR-15x8 magnetron X-band marine radar. The framework combines adaptive intermediate-frequency estimation, conditional residual frequency-offset correction, leakage-window selection, rank-1 singular-value-decomposition reference estimation, pulse-wise phase compensation, matched filtering, and range-tracked time–frequency analysis. Phase-only-correlation timing estimates are retained as diagnostics; the reported products use pass-through timing. Representative Galveston bird and drone collections show range-localized range–Doppler and short-time Fourier-transform products after COR processing, but the scene-dependent workflow is not presented as an aggregate performance benchmark. Second, a kinematic and scattering simulation framework extends a two-segment bird-wing representation to a three-segment model with asymmetric stroke timing, lead–lag motion, body oscillation, optional intermittent-flight gating, aspect-dependent ellipsoidal scattering, and Reynolds-rule flock trajectories. The Herring Gull pa-

rameterization is a simulation design choice rather than a species-validation study. Third, a FAN-integrated Lightweight Hybrid Vision Transformer (FAN-LH-ViT) is evaluated on the separate public DIAT-μSAT benchmark. The model combines multiscale convolutional feature extraction, self-attention, and sine–cosine token projections motivated by periodic micro-motion. The retained checkpoint achieves 98.63% accuracy on a 1,455-sample stratified hold-out subset. The classifier is not trained or evaluated on the Furuno/Galveston collections; consequently, the COR, simulation, and benchmark-classification results are reported as separate contributions rather than as an end-to-end measured-data classification validation.

Chapter 1

Introduction

1.1 Motivation and Research Scope

Small unmanned aerial systems (sUAS) present a difficult radar-classification problem because their radar cross sections and bulk radial velocities can overlap those of birds. Range, received amplitude, and translational Doppler therefore may not provide sufficient class separation under changes in aspect, clutter, and propagation conditions [1]–[3]. Micro-Doppler analysis provides complementary information by resolving phase or frequency modulation produced by target components that move relative to the target body. For the classes considered in this thesis, the relevant micro-motions are primarily flapping wings and rotating propeller blades [4]–[6].

Recovering these modulations from a pulsed radar requires a usable slow-time phase history. The Furuno FAR-15x8 used for the field measurements is a commercial magnetron X-band marine radar. Its pulse-dependent magnetron phase and intermediate-frequency (IF) variation prevent direct application of conventional coherent Doppler processing. Prior coherent-on-receive (COR) systems have shown that a sampled transmit reference can be used to compensate pulse-to-pulse phase variation in magnetron radar data [7], [8]. Building on that principle, this thesis develops and documents a leakage-referenced processing chain for the available Furuno IF collections.

The measured-data contribution is limited to COR processing and analysis of representative range–Doppler and time–frequency products. The Furuno/Galveston collections are not used to train or evaluate the classifier presented later in the thesis. The target-modeling contribution is likewise a simulation study: its bird and drone returns are used to examine how kinematics, scattering assumptions, and viewing geometry affect simulated signatures, not to claim species-level biological validation. The classification contribution is evaluated separately on the public DIAT-μSAT benchmark [9]. Maintaining these dataset and inference boundaries is necessary to distinguish the evidence provided by each part of the thesis.

1.2 Technical Approach

1.2.1 Leakage-Referenced Coherent-on-Receive Processing

The COR framework uses direct transmit leakage in the sampled IF signal as a pulse-specific phase reference. The implemented sequence includes adaptive IF estimation, conditional residual IF-frequency-offset correction, leakage-window detection, rank-1 singular-value-decomposition (SVD) reference estimation, pulse-wise phase rotation, matched filtering, and range tracking. Short-time Fourier-transform (STFT) products are then formed from the range-gated slow-time data. Phase-only-correlation timing estimates are retained as diagnostics, while the reported results use pass-through timing. An additional research extension examines clutter-aware leakage estimation and inter-CPI phase tracking; its performance is assessed through representative scene comparisons rather than a thesis-wide aggregate benchmark.

1.2.2 Kinematic and Scattering Models

The simulation study begins with the two-segment bird-wing representation described by Chen [5]. It introduces a three-segment kinematic chain with asymmetric downstroke and upstroke timing, lead–lag motion, body pitch and heave, optional intermittent-flight gating, and aspect-dependent ellipsoidal scattering. A Herring Gull parameterization is used as a controlled simulation case, with departures from reported morphometrics stated explicitly. The single-bird model is embedded in a Reynolds-rule agent simulation to examine the combined effects of formation geometry, independent wing-beat parameters, and trajectory variation on composite flock signatures [10]. A separate quadcopter point-scatterer model provides a simulated rotor-signature comparison.

1.2.3 Benchmark Classification

Prior studies have applied convolutional neural networks and other supervised classifiers to radar time–frequency representations [11], [12]. The classifier developed here combines a residual squeeze-and-excitation convolutional pyramid with transformer-style self-attention. Fourier Analysis Network (FAN) layers insert sine and cosine projections into the token-processing path, providing an architectural bias toward periodic or quasi-periodic structure [13], [14]. The resulting FAN-integrated Lightweight Hybrid Vision Transformer (FAN-LH-ViT) is trained and evaluated using a stratified hold-out split of DIAT-μSAT. Because the benchmark preprocessing and split are not reproduced identically from all prior studies, the reported accuracy is not treated as a controlled state-of-the-art comparison.

1.3 Thesis Contributions

The principal contributions are:

1. **Furuno IF acquisition and COR processing:** documentation of the FAR-15x8 IF-stage acquisition configuration and implementation of a leakage-referenced processing chain that produces pulse-phase-compensated, matched-filtered data for range–Doppler and time–frequency analysis.

2. **Range-tracked analysis of representative field collections:** application of the maintained COR workflow to selected Galveston bird and drone scenes, with explicit separation between applied corrections, diagnostic-only processing, interactive range selection, and scene-dependent results.

3. **Kinematic and scattering simulation of aerial micro-Doppler:** formulation of a three-segment bird-wing model, Reynolds-rule flock trajectories, aspect-angle analysis, and a quadcopter comparison for examining how simulated target motion and scattering parameters affect Doppler signatures.

4. **Compact benchmark classifier:** development of the FAN-LH-ViT architecture and evaluation of its retained checkpoint on a 1,455-sample stratified DIAT-μSAT hold-out subset, where it attains 98.63% accuracy with 2,375,918 trainable parameters.

These contributions are related through the micro-Doppler classification problem but are not evaluated as a single end-to-end Furuno classification system.

1.4 Organization of the Thesis

Chapter 2 reviews the radar signal model, range and Doppler processing, coherent processing intervals, and the phase-reference problem posed by a magnetron transmitter. Chapter 3 develops the micro-Doppler formulation and the bird, flock, aspect-angle, and drone simulations. Chapter 4 formulates supervised spectrogram classification and reviews the machine-learning methods most relevant to the classifier study. Chapter 5 describes the Furuno acquisition context, COR signal model, implemented processing sequence, and representative field products. Chapter 6 presents the FAN-LH-ViT architecture, DIAT-μSAT preprocessing, training protocol, and hold-out results. Chapter 7 summarizes the supported findings, limitations, and future work.

Chapter 2

Multifunction Radar

2.1 Radar Signal Model and Measurement Variables

Radar estimates target parameters from electromagnetic energy scattered toward the receiver. The measured delay, complex amplitude, and pulse-to-pulse phase history provide information about range, scattering strength, and radial motion [2], [15]. X-band systems operate nominally from 8 GHz to 12 GHz; the corresponding wavelength permits compact antennas and produces Doppler shifts suitable for resolving wing and rotor micro-motion.

2.1.1 The Transmitted Radar Signal

A real passband radar waveform can be written as [2], [16]

$$x(t) = a(t)\cos\left[2\pi f_0 t + \theta(t)\right] \tag{2.1}$$

where

- $x(t)$ is the real passband waveform.
- $a(t)$ is the amplitude envelope (modulation), which may be constant for continuous-wave radar or time-varying (e.g., a rectangular pulse) for pulsed

radar.

- f_0 is the carrier frequency, which determines the radar's operating band.
- $\theta(t)$ is the phase modulation, when present.

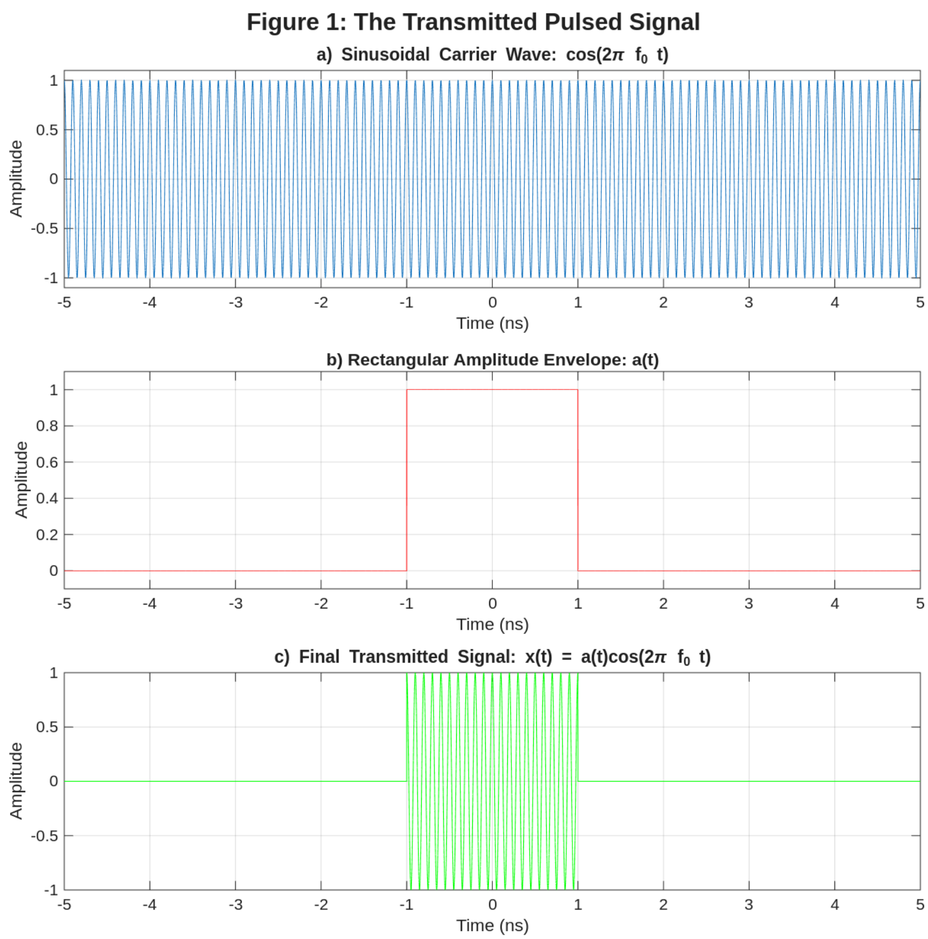


Figure 2.1: **Model of the Transmitted and Received Signal.** (a) The blue pulse represents the carrier wave of $x(t)$. (b) The red pulse is a rectangular amplitude envelope of $x(t)$. (c) The green pulse is the instantaneous amplitude of the transmitted signal at time t.

The real passband signal $x(t)$ is obtained by modulating a baseband waveform onto the carrier f_0. The choice of $a(t)$ and $\theta(t)$ is central to radar waveform design, as it affects range and velocity measurement, target resolution, and robustness to interference. In analysis and implementation, it is often more convenient to work with the complex baseband (complex envelope) of the transmitted signal, defined as [2]:

$$\tilde{x}(t) = a(t)e^{j\theta(t)} \tag{2.2}$$

with the real signal recovered as

$$x(t) = \Re\left\{\tilde{x}(t)e^{j2\pi f_0 t}\right\}. \tag{2.3}$$

This representation simplifies mathematical analysis and reflects how signals are processed in modern digital radar receivers. Waveform design typically focuses on $\tilde{x}(t)$; the carrier f_0 mainly influences antenna characteristics and propagation. In systems employing high-power amplifiers (HPAs), such as traveling-wave tubes (TWTs) or solid-state power amplifiers (SSPAs), $a(t)$ is often constrained to be rectangular (constant during transmission and zero otherwise) to operate in saturation and maximize power efficiency [2]. In such cases, $\theta(t)$ becomes the primary degree of freedom for tailoring the waveform to specific operational requirements.

2.1.2 The Received Radar Signal

After transmission, $x(t)$ propagates through the medium (typically free space or the atmosphere) at approximately the speed of light $c \approx 3 \times 10^8$ m/s. When the wave encounters a target, a portion of its energy is scattered; the receive antenna collects the fraction that returns toward the radar. For a single point scatterer and neglecting clutter and interference, a convenient baseband model for the received signal is [2]

$$y(t) = \alpha x(t - \tau_d) + n(t), \tag{2.4}$$

where α is a complex scaling factor accounting for two-way propagation losses, target radar cross section (RCS), antenna gains, and phase shifts; τ_d is the two-way propagation delay; and $n(t)$ is typically modeled as additive white Gaussian noise (AWGN). Using the complex-envelope representation and explicitly including a Doppler shift f_D, the received signal from a point target can be written as

$$\tilde{y}(t) = \hat{a}(t - \tau_d) \exp\left\{j\left[2\pi f_D(t - \tau_d) + \theta(t - \tau_d) + \phi_s(t)\right]\right\} + \tilde{n}(t), \tag{2.5}$$

where $\hat{a}(\cdot)$ is the received amplitude (including range, RCS, and antenna effects), $\theta(\cdot)$ represents system-related phase terms, $\phi_s(t)$ captures additional target- or channel-induced phase modulation (time-varying for complex targets with micromotions and a source of micro-Doppler), and $\tilde{n}(t)$ is the complex-envelope noise. In realistic environments, the received signal is a superposition of returns from multiple targets, environmental clutter, and possible interference. Separating target echoes from these unwanted components is a central goal of radar signal processing and directly impacts the quality of data used in subsequent classification stages.

2.1.3 Range Measurement and Resolution

A primary function of radar is to estimate target range. For a monostatic radar, with co-located transmit and receive antennas, the two-way delay τ_d is proportional to the range R:

$$R = \frac{c\tau_d}{2}, \tag{2.6}$$

where c is the speed of light. The ability to distinguish two closely spaced targets in range is characterized by the range resolution ΔR. For a waveform with effective (occupied) bandwidth B, the fundamental limit on ΔR is [2], [17]

$$\Delta R = \frac{c}{2B}. \tag{2.7}$$

This follows from the minimum resolvable delay difference $\Delta\tau \approx 1/B$ and $\Delta R = c\Delta\tau/2$. For simple unmodulated pulses of duration T_p, the effective bandwidth is on the order of $B \approx 1/T_p$, so improving range resolution requires shortening T_p. The pulse energy is approximately

$$E_p \approx P_t T_p, \tag{2.8}$$

where P_t is the transmitted power. Thus, reducing T_p while holding P_t fixed reduces E_p, which can decrease the maximum detection range. Waveform techniques such as pulse compression are therefore used to obtain large effective bandwidths while retaining longer pulse durations and adequate transmitted energy. In digital radars, the received analog signal is sampled by an ADC after each transmitted pulse. Samples are taken at intervals of T_s starting at time $t_{\min}$ after transmission. The l-th sample corresponds to the delay

$$\tau_d = t_{\min} + lT_s, \quad l = 0, 1, \ldots, L-1, \tag{2.9}$$

which is converted to range using (2.6). This defines a discrete grid of range values, where each index l is associated with a *range bin* (fast-time sample). The spacing between adjacent bins,

$$\Delta R_{\text{bin}} = \frac{cT_s}{2}, \tag{2.10}$$

is typically chosen to be consistent with or finer than the intrinsic range resolution ΔR. These fast-time samples form the range dimension of the data that will be combined across pulses to form range–Doppler and related representations.

2.1.4 Doppler Frequency Shift and Velocity Measurement

If a target moves with radial velocity v_r relative to the radar, the received frequency shifts from f_0 by an amount known as the Doppler shift. For $v_r \ll c$, the shift is [2], [18]

$$f_D = \frac{2v_r}{\lambda}, \tag{2.11}$$

where $\lambda = c/f_0$ is the radar wavelength. A more exact relativistic expression is

$$f_D = f_0 \left(\sqrt{\frac{c + v_r}{c - v_r}} - 1 \right), \tag{2.12}$$

which reduces to (2.11) for $v_r \ll c$. Measuring f_D enables estimation of the radial velocity as

$$v_r = \frac{f_D \lambda}{2}. \tag{2.13}$$

In pulsed radar, the Doppler shift appears as a systematic phase progression of the complex baseband signal from pulse to pulse. If $s[m]$ denotes the complex return from a given range bin at pulse index m (slow time), then for a single moving point target

$$s[m] \approx A e^{j(2\pi f_D m T_{\mathrm{PRI}} + \phi_0)} + w[m], \tag{2.14}$$

where A is a complex amplitude, T_{PRI} is the pulse repetition interval, ϕ_0 is an initial phase, and $w[m]$ denotes noise and other disturbances. The Doppler frequency f_D can be estimated via a discrete Fourier transform (DFT) across m, which underlies range–Doppler map formation in coherent pulsed radar. The ability to measure Doppler shift allows radar systems to distinguish moving targets from stationary clutter (e.g., via Moving Target Indication, MTI) and to estimate target velocities. In later chapters, Doppler and micro-Doppler effects are exploited to form time–frequency signatures that feed the deep learning classification pipeline.

2.1.5 Signal Strength and Radar Cross Section (RCS)

The received signal strength depends on transmitted power, antenna characteristics, propagation loss, range, and target scattering. Radar cross section (RCS), σ (in m^2), is the equivalent area that relates incident power density to power scattered in the radar direction [15]. Combining the target-scattering definition with the transmit and receive aperture factors gives the radar range equation in Section 2.1.6. The RCS generally differs from physical cross-sectional area and depends on target size, shape, material, orientation, operating frequency, and polarization. It can vary significantly with aspect angle and time, especially for complex targets exhibiting micromotions (e.g., rotating blades or flapping wings). For simple canonical shapes (spheres, plates, cylinders), RCS can be estimated analytically. For complex targets (e.g., aircraft, birds), it is usually obtained via measurements or high-fidelity EM simulations. Statistical RCS models, such as the Swerling cases, describe fluctuations in detection performance. For example, Swerling Case 1 assumes the RCS is constant during a single scan (dwell) but varies independently from scan to scan, with an exponential probability density function for σ (or Rayleigh-distributed voltage amplitude) [2]. These models directly influence predicted detection probabilities through the radar range equation and SNR calculations.

2.1.6 The Radar Range Equation

The radar range equation links the average received power to transmitter, receiver, antenna, target, and propagation characteristics. It underpins detection-range predictions and key radar design trade-offs. For a monostatic radar operating in free space, the average received power P_r from a point target with RCS σ at range R is [2], [15]

$$P_r = \frac{P_t G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4 L}, \tag{2.15}$$

where P_t is transmitted power, G_t and G_r are transmit and receive antenna gains (often $G_t = G_r = G$), λ is the wavelength, and L represents aggregate system losses (e.g., radome, ohmic, processing). The characteristic $1/R^4$ dependence arises from $1/R^2$ spreading from transmitter to target and another $1/R^2$ from target back to receiver. Target detectability is governed by the signal-to-noise ratio (SNR) at the receiver output. If the receiver has noise figure F, noise-equivalent bandwidth B_n, and operates at temperature T_0, the noise power is

$$P_n = kT_0 F B_n, \tag{2.16}$$

where k is Boltzmann's constant. The resulting SNR is

$$\mathrm{SNR} = \frac{P_r}{P_n} = \frac{P_t G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4 L k T_0 F B_n}. \tag{2.17}$$

For pulsed radar, it is convenient to use energy quantities. If the pulse energy is $E_p = P_t T_p$ and the matched-filter receiver integrates over the pulse duration, the single-pulse SNR is

$$\mathrm{SNR}_{\mathrm{pulse}} = \frac{E_p G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4 L k T_0 F}, \tag{2.18}$$

where the matched filter accounts for the noise bandwidth. Coherent integration of M identical pulses over a coherent processing interval (CPI) ideally improves SNR by a factor of M:

$$\mathrm{SNR}_{\mathrm{CPI}} = M \, \mathrm{SNR}_{\mathrm{pulse}}, \tag{2.19}$$

provided that target parameters remain essentially constant over the CPI. Rearranging (2.15) or the SNR expression to solve for R yields the *maximum detection range* for a specified minimum SNR. This quantity is central to radar link-budget analysis and to characterizing the performance envelope of the non-coherent pulsed radar system considered in this thesis.

2.2 Pulsed Radar Systems

Pulsed radar systems transmit short bursts (pulses) of EM energy and then listen for echoes from targets during the quiet interval between pulses. By measuring the round-trip time of the pulse, the radar determines target range. Pulse characteristics and timing are fundamental to radar performance and its inherent range/velocity ambiguities [2], [15].

2.2.1 The Simple Pulse Waveform

The simplest pulsed-radar waveform is an unmodulated rectangular RF pulse. The complex envelope $\tilde{x}(t)$ for a single pulse of duration T_p and unit amplitude is

$$\tilde{x}(t) = \mathrm{rect}\left(\frac{t}{T_p}\right) = \begin{cases} 1, & 0 \le t \le T_p \\ 0, & \text{otherwise} \end{cases} \tag{2.20}$$

with corresponding real transmitted signal

$$x(t) = \mathrm{rect}\left(\frac{t}{T_p}\right)\cos(2\pi f_0 t). \tag{2.21}$$

The energy in a single pulse of amplitude A is $E_p = A^2 T_p$ (assuming unit impedance). The pulse bandwidth B is approximately inversely proportional to its duration, $B \approx 1/T_p$. From (2.7), the resulting range resolution is

$$\Delta R \approx \frac{cT_p}{2}. \tag{2.22}$$

This reveals a fundamental conflict for simple pulsed radars: fine range resolution requires short T_p, but shorter pulses reduce energy for fixed peak power, decreasing maximum detection range. This motivates the use of pulse-compression waveforms that combine the energy of a long pulse with the resolution of a short pulse.

2.2.2 Pulse Repetition Interval and Frequency

In pulsed radar, pulses are usually transmitted periodically. The time between the start of consecutive pulses is the pulse repetition interval (PRI), T_{PRI}, and its reciprocal is the pulse repetition frequency (PRF), f_{PRF}:

$$f_{\text{PRF}} = \frac{1}{T_{PRI}}. \tag{2.23}$$

The PRF (or PRI) is a critical design parameter controlling the radar's unambiguous range and Doppler-velocity coverage [2], [19].

Unambiguous Range (R_{ua})

The maximum range from which an echo can return before the next pulse is transmitted is the maximum unambiguous range, R_{ua}. This occurs when the round-trip time equals the PRI:

$$\tau_d = T_{PRI} = \frac{2R_{ua}}{c}. \tag{2.24}$$

Thus,

$$R_{ua} = \frac{cT_{PRI}}{2} = \frac{c}{2f_{\mathrm{PRF}}}. \tag{2.25}$$

Targets beyond R_{ua} produce echoes that arrive after one or more subsequent pulses have been transmitted, causing their apparent range to be folded into $[0, R_{ua}]$. If a true range is $R_{true} = kR_{ua} + R_{apparent}$, k integer and $0 \leq R_{apparent} < R_{ua}$, the radar measures $R_{apparent}$. Techniques such as PRF jittering or multiple-PRF processing can resolve these ambiguities at the cost of added complexity.

Unambiguous Doppler Velocity (v_{ua})

The PRF also limits unambiguous Doppler measurement. The Doppler-induced phase is effectively sampled once per pulse at rate f_{PRF}. By the Nyquist-Shannon theorem, the maximum unambiguous Doppler frequency is $\pm f_{\mathrm{PRF}}/2$:

$$|f_D|_{\max} = \frac{f_{\mathrm{PRF}}}{2}. \tag{2.26}$$

Using (2.11), the maximum unambiguous radial velocity is

$$|v_r|_{ua} = \frac{f_{\mathrm{PRF}}\lambda}{4}. \tag{2.27}$$

Targets with radial speed above this limit are measured at aliased Doppler frequencies inside the principal interval. Staggered PRFs and related techniques can help resolve Doppler ambiguities.

The Range–Doppler Dilemma

Equations (2.25) and (2.27) imply a fundamental trade-off in pulsed radar design, often called the range–Doppler dilemma. Multiplying R_{ua} and $|v_r|_{ua}$ gives

$$R_{ua}|v_r|_{ua} = \left(\frac{c}{2f_{\mathrm{PRF}}}\right)\left(\frac{f_{\mathrm{PRF}}\lambda}{4}\right) = \frac{c\lambda}{8}. \tag{2.28}$$

For a given λ, this product is constant: increasing PRF to improve unambiguous velocity reduces unambiguous range, and vice versa.

- **High PRF mode:** Good unambiguous velocity coverage but significant range ambiguities; used when velocity is primary and range ambiguities can be tolerated or resolved (e.g., airborne intercept radars in look-down modes).
- **Low PRF mode:** Good unambiguous range but strong Doppler ambiguities (low blind speeds); used in long-range search and surveillance where initial detection and ranging are prioritized.
- **Medium PRF mode:** A compromise with both range and Doppler ambiguities present; requires multi-PRF ambiguity-resolution processing.

The choice of PRF is therefore strongly application- and environment-dependent.

2.2.3 Coherent Processing Intervals (CPI)

To accurately measure Doppler, radar typically transmits and processes a sequence of M coherent pulses, known as a coherent processing interval (CPI) or dwell, of duration $T_{\mathrm{CPI}} = MT_{PRI}$ [2], [19]. "Coherent" implies that the phase relationship between transmitted pulses is known and stable, and the receiver maintains phase coherence over the CPI. This permits energy integration from multiple pulses and extraction of Doppler information. Within a CPI, data may be viewed as a two-dimensional matrix:

- **Fast-time:** Samples within a single PRI, corresponding to echoes from different ranges. If L range bins are collected per pulse, this dimension has L samples and determines range-bin spacing via the ADC sampling rate.
- **Slow-time:** Samples from the same range bin across the M pulses in the CPI. The slow-time sampling rate is f_{PRF}, and variations are due to target motion (Doppler) and other dynamics such as micro-Doppler or RCS fluctuations.

For a stationary target, the complex amplitude in a given range bin is ideally constant over M pulses (ignoring noise and RCS fluctuations). For a moving target, the phase evolves linearly in slow-time at frequency f_D. Processing these M slow-time samples (e.g., via a DFT) allows estimation of f_D. The Doppler resolution Δf_D, i.e., the ability to distinguish closely spaced Doppler frequencies, is inversely proportional to CPI duration:

$$\Delta f_D = \frac{1}{T_{\text{CPI}}} = \frac{1}{MT_{PRI}} = \frac{f_{\text{PRF}}}{M}. \tag{2.29}$$

A longer CPI (larger M) yields finer Doppler resolution, but requires that the target remain within the same range bin and maintain approximately constant velocity and scattering characteristics over T_{CPI}, which may not hold for highly maneuvering or rapidly micro-moving targets. Coherent integration over M pulses also improves SNR. If signals add coherently and noise adds incoherently, the SNR gain relative to a single pulse is up to M:

$$\text{SNR}_{\text{CPI}} = M\,\text{SNR}_{\text{pulse}}. \tag{2.30}$$

This processing gain improves detection of low-SNR targets.

2.2.4 Magnetron Pulse Phase and Coherent-on-Receive Processing

The preceding Doppler formulation assumes that the phase reference is consistent from pulse to pulse. A magnetron transmitter does not provide this condition directly: each transmitted pulse has an uncontrolled phase origin, and the sampled IF can also exhibit pulse-dependent frequency variation. If the received complex sample at a fixed range is multiplied by an unknown factor $e^{j\phi_m}$ on pulse m, a slow-time DFT contains the desired target phase progression together with the sequence ϕ_m. The resulting spectrum does not represent target Doppler alone.

Coherent-on-receive processing estimates a pulse-specific reference from a sampled copy of the transmitted waveform or from direct transmit leakage [7], [8]. Because the leakage and target echoes share the same pulse-dependent phase origin, rotation by the conjugate leakage phase removes the common term while retaining the target-to-reference propagation-phase difference. Residual IF-frequency error must be estimated separately because it produces a fast-time phase slope and a delay-dependent phase error after downconversion. Chapter 5 develops the pulse-indexed analytic IF model and documents the correction sequence used for the Furuno measurements.

This distinction also bounds the classification study. The Furuno chapters analyze COR-processed field products, whereas Chapter 6 trains and evaluates the classifier on the separate DIAT-μSAT benchmark. Envelope-only Furuno matrices are not used as inputs to the reported classifier.

2.3 Chapter Summary

This chapter defined the passband and complex-envelope signal models, range and Doppler measurements, the monostatic radar range equation, pulsed-radar ambiguities, and coherent processing intervals. It then identified the uncontrolled pulse phase and IF variation of a magnetron radar as the specific signal-processing problem addressed by leakage-referenced COR in Chapter 5.

Chapter 3

Micro-Doppler Signature

3.1 Fundamentals of the Micro-Doppler Effect

3.1.1 Overview

Micro-Doppler denotes time-varying Doppler modulation produced by motion of scattering components relative to a target reference point. Vibration, rotation, oscillation, and articulation can produce sidebands or time-varying ridges around the translational Doppler component. The resulting signature depends on component kinematics, scattering amplitude, wavelength, pulse repetition frequency, and viewing geometry [4].

The same phase-modulation principle applies to laser and microwave radar [5]. At microwave wavelengths, rotating blades, flapping wings, and articulated limbs can produce component velocities large enough to generate resolvable Doppler modulation when the radar provides adequate phase stability, sampling rate, and signal-to-clutter ratio.

Extraction and interpretation of micro-Doppler signatures have become central to radar-based target discrimination. Applications include human activity recognition, drone detection and classification, wind-turbine clutter identification, and, as emphasized in this chapter, distinguishing bird targets from small unmanned aerial systems (sUAS) [6].

3.1.2 Physical and Mathematical Description

Consider a monostatic radar transmitting a continuous-wave (CW) signal at carrier frequency f_0. Define radial velocity v_r as positive for a closing target. The Doppler shift is

$$f_D = \frac{2v_r}{\lambda}, \tag{3.1}$$

where $\lambda = c/f_0$. When component motion is present, the closing radial velocity of a scattering point can be decomposed into translational and component-motion terms:

$$v_r(t) = v_{r,0} + v_{r,\mathrm{m}}(t), \tag{3.2}$$

giving a time-varying Doppler frequency

$$f_D(t) = \frac{2}{\lambda}\Big[v_{r,0} + v_{r,\mathrm{m}}(t)\Big]. \tag{3.3}$$

To formalize the geometry, two coordinate systems are introduced: a global system (X, Y, Z) in which the radar is at the origin Q, and a body-fixed system (x, y, z) that translates and rotates with the target [5]. A point scatterer P initially located at position $\mathbf{r}_0$ in the body frame has a global position at time t of

$$\mathbf{R}_P(t) = \mathbf{R}_0 + \mathbf{v}t + \mathfrak{R}(t)\,\mathbf{r}_0, \tag{3.4}$$

where $\mathbf{R}_0$ is the initial position of the body origin relative to the radar, $\mathbf{v}$ is the translational velocity, and $\mathfrak{R}(t)$ is the body-to-global rotation matrix. Using the baseband convention adopted in the simulation, the two-way propagation phase is

$$\Phi_P(t) = -\frac{4\pi}{\lambda}\,\|\mathbf{R}_P(t)\|. \tag{3.5}$$

The instantaneous Doppler frequency is

$$f_D = \frac{1}{2\pi}\frac{d\Phi_P}{dt} = -\frac{2}{\lambda}\left[\mathbf{v} + \frac{d}{dt}\Big(\mathfrak{R}(t)\,\mathbf{r}_0\Big)\right]^T \hat{\mathbf{n}}, \tag{3.6}$$

where $\hat{\mathbf{n}}$ points from the radar to the scatterer. When the target has angular velocity $\boldsymbol{\Omega}$, $\frac{d}{dt}(\mathfrak{R}\mathbf{r}_0) = \boldsymbol{\Omega} \times \mathbf{r}$, with $\mathbf{r} = \mathfrak{R}\mathbf{r}_0$. Under the far-field approximation, the Doppler frequency separates as

$$f_D \approx \underbrace{-\frac{2}{\lambda}\,\mathbf{v}\cdot\hat{\mathbf{n}}}_{f_{\mathrm{trans}}} + \underbrace{-\frac{2}{\lambda}\,(\boldsymbol{\Omega}\times\mathbf{r})\cdot\hat{\mathbf{n}}}_{f_{\mathrm{mD}}}. \tag{3.7}$$

The first term is the translational Doppler; the second is the micro-Doppler component arising from rotation or oscillation of body parts.

For a vibrating scatterer with sinusoidal displacement $D_v(t) = D_{\max}\sin(2\pi f_v t)$, the micro-Doppler shift reduces to

$$f_{\mathrm{mD}}(t) = \frac{4\pi f_0\, f_v\, D_{\max}}{c}\cos(2\pi f_v t)\, G(\alpha, \beta, \alpha_P, \beta_P), \tag{3.8}$$

where $G(\cdot)$ represents the geometric projection of the vibration direction onto the radar LOS. The maximum micro-Doppler bandwidth is therefore proportional to the product $f_v D_{\max}$, and the signature of such a scatterer is a sinusoidal course in the time–frequency plane.

3.2 Time–Frequency Analysis of Micro-Doppler Signatures

3.2.1 Short-Time Fourier Transform

Because micro-Doppler modulations are time varying, a joint time–frequency representation is required. A commonly used representation is the short-time Fourier transform (STFT). For a discrete signal $x[n]$ and a window function $w[n]$ of length M:

$$\mathrm{STFT}[m,k] = \sum_{n=0}^{M-1} x[n + m\,H]\,w[n]\,e^{-j2\pi kn/N_{\mathrm{FFT}}} \tag{3.9}$$

where H is the hop size and N_{FFT} is the FFT length. The spectrogram is $|\,\mathrm{STFT}[m,k]|^2$. The simulation figures in this chapter use a Kaiser window ($\beta = 10$, $M = 256$), 250-sample overlap, and $N_{\mathrm{FFT}} = 512$. Chapter 5 and Chapter 6 state their own STFT parameters because the measured-data and benchmark-classification pipelines are different.

Window length controls the usual time–frequency resolution trade-off: a shorter window localizes transients more precisely but broadens narrowband components, whereas a longer window provides finer frequency sampling over a longer local interval.

3.2.2 Alternative Time–Frequency Methods

Beyond the STFT, several alternative representations have been applied to micro-Doppler analysis:

- **Empirical Mode Decomposition (EMD) and the Hilbert–Huang Transform (HHT):** EMD adaptively decomposes a signal into Intrinsic Mode Functions (IMFs), each of which admits a well-defined instantaneous frequency via the Hilbert transform [20]. The method is data-driven and imposes no fixed basis, making it suitable for nonlinear and nonstationary micro-Doppler signals.

- **Wavelet Transform:** Multi-resolution analysis using wavelets can provide better frequency resolution at low frequencies and better time resolution at high frequencies, complementing the fixed-resolution STFT.
- **Wigner–Ville Distribution:** Provides optimal time–frequency concentration for mono-component signals but suffers from cross-term interference for multi-component returns, as is typical of bird wing micro-Doppler.

The simulation results in this chapter use the STFT because it provides a common representation for comparing multi-scatterer bird and drone returns. The classifier in Chapter 6 independently generates STFT magnitude images from DIAT-μSAT signals; it does not use the simulated spectrograms from this chapter.

3.3 Bird Micro-Doppler Phenomenology

A flying bird generates micro-Doppler signatures through the periodic flapping of its wings. The body itself translates at a roughly constant cruising speed, producing a near-constant Doppler line, while the oscillating wing segments impose time-varying Doppler modulations whose frequency extent, harmonic content, and temporal asymmetry depend primarily on the underlying wing kinematics [5].

3.3.1 Wing-Beat Characteristics

Bruderer et al. [21] compiled wing-beat frequencies for over 150 bird species, reporting typical values of 2 Hz to 15 Hz depending on body mass and wingspan. The empirical relationship derived by Pennycuick [22] predicts wing-beat frequency as:

$$f_w = m^{0.375}\, b_w^{-0.958}\, S^{-0.333} \left(\frac{g}{\rho}\right)^{0.5} \tag{3.10}$$

where m is body mass, b_w is wingspan, S is wing area, g is gravitational acceleration, and ρ is air density. Tracking-radar measurements provide an empirical point of comparison for this relationship [23].

A key observation from cine camera and radar studies is that the downstroke typically occupies approximately 60% of the wing-beat period, with the upstroke taking the remaining 40% [24], [25]. This temporal asymmetry has direct consequences for the micro-Doppler signature: the ridges produced during the downstroke and upstroke have different dwell times, leading to asymmetric harmonic amplitudes. Many species also exhibit *intermittent flight*, alternating between flapping bursts (0.1 s to 0.4 s) and passive (ballistic or gliding) pauses (0.1 s to 0.6 s) [21].

3.3.2 Avian Wing Morphology

Anatomically, a bird's wing comprises three articulated segments:

1. **Arm (humerus):** The innermost segment, connecting the shoulder joint to the elbow.
2. **Forearm (radius/ulna):** The middle segment, bearing the secondary flight feathers, extending from the elbow to the wrist.
3. **Hand (manus):** The outermost segment, bearing the primary flight feathers, extending from wrist to wingtip.

The model assigns segment-dependent flapping (elevation angle ξ_k), lead–lag or sweep (azimuth angle ζ_k), and twist (pronation/supination angle θ_k) [24], [26]. These degrees of freedom provide a kinematic basis for studying how articulated motion changes the simulated Doppler trajectories.

3.3.3 Impact on Radar Returns

The micro-Doppler bandwidth of a bird is governed by the radial velocity of the outermost scattering point—the wingtip. For a bird with half-wingspan b (all segments combined), flapping at frequency f_w through an angular excursion $\Delta\xi$, the peak wingtip velocity is approximately $v_{\text{tip}} \approx 2\pi f_w\, b\, \sin(\Delta\xi/2)$. The maximum micro-Doppler extent is then

$$f_{\text{mD,max}} = \frac{2\, v_{\text{tip}}}{\lambda}. \tag{3.11}$$

At X-band ($\lambda \approx 3.2\,\text{cm}$), a medium-sized bird with $b = 0.5\,\text{m}$, $f_w = 4\,\text{Hz}$, and $\Delta\xi = 70°$ produces $v_{\text{tip}} \approx 7.3\,\text{m/s}$ and $f_{\text{mD,max}} \approx 460\,\text{Hz}$—well within the unambiguous Doppler range of a radar with PRF $> 1\,\text{kHz}$.

The bird's body, being comparatively large and rigid, dominates the return at low Doppler offsets. The wing segments contribute progressively weaker but progressively wider Doppler modulations, with the hand/primary segment generating the highest instantaneous velocities and therefore the broadest Doppler bandwidth.

3.4 Chen's Two-Jointed Arm Model

3.4.1 Model Description

Chen summarizes a two-jointed arm model that is used here as the simulation baseline [5]. The model represents each wing as two rigid segments of equal length $L_1 = L_2 = 0.5\,\text{m}$ connected at an elbow joint. The flap angles are driven by a symmetric cosine:

$$\xi_k(t) = \xi_{k,0} + A_k \cos(2\pi f_w t), \quad k \in \{1, 2\} \tag{3.12}$$

where A_k is the amplitude and $\xi_{k,0}$ is the offset angle. The forearm also maintains a constant twist angle, C_2. Five ellipsoidal scatterers are placed at the body center and at the midpoints of the four wing segments (two per side). The radar return is the coherent sum of the five range-gated echoes.

3.4.2 Identified Limitations

The baseline omits several kinematic and scattering elements considered in the avian-flight literature. Table 3.1 lists those simplifications and the corresponding expected changes in the simulated spectrogram. These entries motivate model extensions; they are not a validation ranking against measured birds.

Table 3.1: Limitations of the Chen two-jointed arm model.

#	Limitation	Physical basis	Spectrogram impact
1	Symmetric cosine stroke	Reported downstroke fraction $\approx 60\%$ [24], [25]	Equal modeled ridge dwell times
2	Only 2 wing segments	Bird wing has arm + forearm + hand [26]	Missing distal hand kinematics; narrower bandwidth
3	No lead-lag / sweep	All segments exhibit fore-aft sweep [24]	Ridge skew absent
4	Single constant twist	Twist differs between strokes [24]	No stroke-dependent orientation term
5	No body oscillation	Body pitches and heaves each cycle [26]	No body-line FM
6	No intermittent flight	Common in passerines [21]	No burst-pause gaps
7	5 lumped scatterers	3 segments yield 7 scatterers	Coarser spatial sampling

3.5 Three-Segment Wing Model

This section introduces a three-segment bird micro-Doppler model. Six kinematic and temporal modifications address the seven baseline simplifications in Table 3.1; the seventh item, scatterer count, follows from adding the distal segment. Figure 3.1 compares the model geometries.

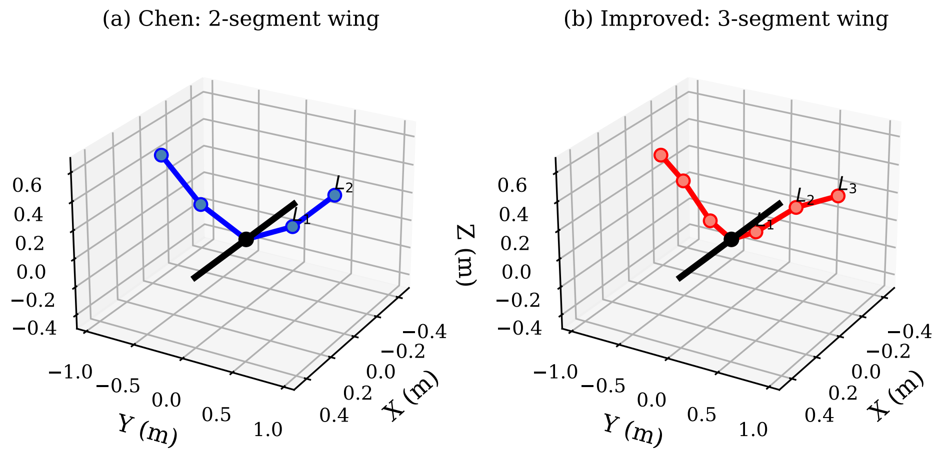


Figure 3.1: Comparison of (a) the Chen two-segment wing model and (b) the three-segment wing model with lead–lag sweep. Each circle denotes a scatterer center.

3.5.1 Asymmetric Stroke Function

The symmetric cosine of Eq. (3.12) is replaced by an asymmetric piecewise-cosine function that independently controls the downstroke and upstroke durations [5]. Let $\tau = (f_w\, t) \bmod 1$ be the normalized phase within each beat cycle and let η_d denote the downstroke fraction. The stroke function is

$$g_f(\tau;\, \eta_d) = \begin{cases} \cos\left(\dfrac{\pi\,\tau}{\eta_d}\right), & 0 \le \tau < \eta_d \\[2ex] -\cos\left(\dfrac{\pi\,(\tau - \eta_d)}{1 - \eta_d}\right), & \eta_d \le \tau < 1 \end{cases} \tag{3.13}$$

where $g_f = +1$ at the top of the stroke and $g_f = -1$ at the bottom. The function is plotted in Fig. 3.2 for $\eta_d = 0.60$ (60% downstroke and 40% upstroke), based on the cited kinematic literature [24], [25]. The unequal durations change the modeled angular velocity on the two half-cycles; aerodynamic force is not computed.

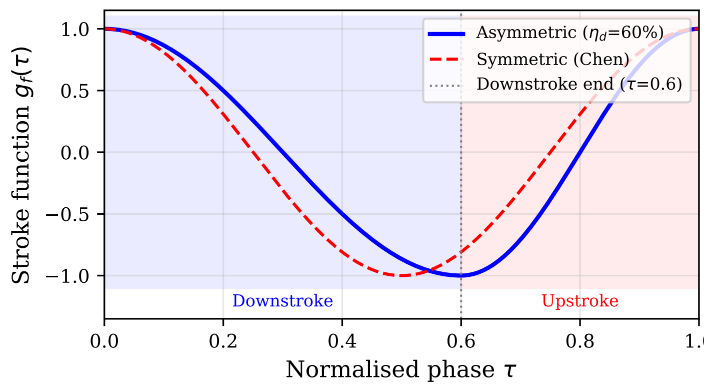


Figure 3.2: Comparison of the asymmetric piecewise-cosine stroke function ($\eta_d = 0.60$) used in the three-segment model with the symmetric cosine used in the baseline. The shaded regions indicate the downstroke and upstroke phases.

3.5.2 Three-Segment Wing Forward Kinematics

Each wing is modeled as a kinematic chain of three rigid segments—arm (L_1), forearm (L_2), and hand (L_3)—articulated at the shoulder, elbow, and wrist joints. The geometric model is illustrated in Fig. 3.3, and the kinematic chain is diagrammed in Fig. 3.4.

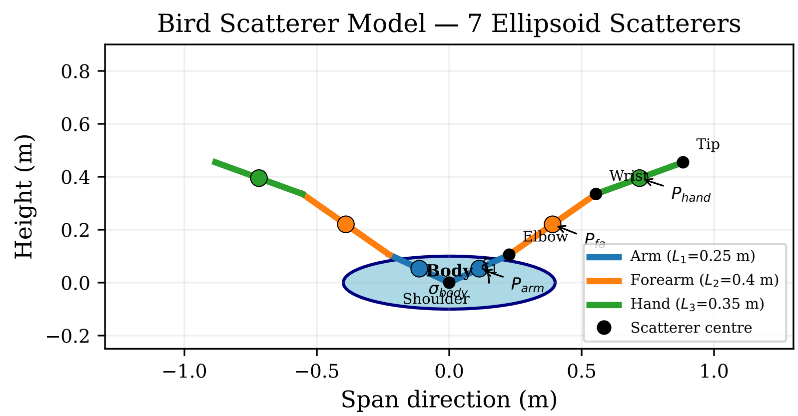


Figure 3.3: Bird scatterer geometry showing the seven ellipsoidal scatterers: one for the body and three per wing (arm, forearm, hand). Scatterer centers are placed at the midpoints of each segment.

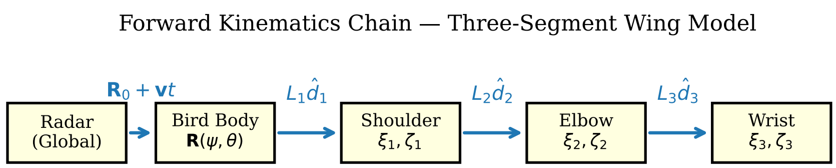


Figure 3.4: Forward kinematics chain for the three-segment wing model. Each joint introduces a direction vector $\hat{d}_k$ parameterised by flap (ξ_k) and sweep (ζ_k) angles.

Each segment k ($k = 1, 2, 3$ for arm, forearm, hand) has two primary angular degrees of freedom:

- **Flap angle** $\xi_k(t)$: elevation in the plane perpendicular to the flight axis.
- **Lead-lag (sweep) angle** $\zeta_k(t)$: fore-aft displacement in the horizontal plane.

The flap angles are driven by the asymmetric stroke function:

$$\xi_k(t) = \xi_{k,0} + A_k\, g_f(\tau;\, \eta_d), \quad k = 1, 2, 3 \tag{3.14}$$

while the lead-lag angles are modeled as cosine oscillations with segment-dependent phase offsets:

$$\zeta_k(t) = \zeta_{k,0} + \hat{\zeta}_k \cos(2\pi f_w t + \phi_{\zeta,k}). \tag{3.15}$$

Given these angles, the direction vector of segment k in the body frame (with the span axis along $+y$) is

$$\hat{\mathbf{d}}_k = \begin{bmatrix} -\sin\zeta_k \cos\xi_k \\ \cos\zeta_k \cos\xi_k \\ \sin\xi_k \end{bmatrix}. \tag{3.16}$$

At rest ($\xi_k = 0$, $\zeta_k = 0$) this gives $\hat{\mathbf{d}} = [0,\ 1,\ 0]^T$, i.e. the segment points along the span. A positive ξ_k raises the segment above the horizontal ($z > 0$), while a positive ζ_k sweeps the segment aft ($x < 0$).

The joint positions are obtained by forward summation along the kinematic chain:

$$\mathbf{p}_{\text{elbow}} = L_1\, \hat{\mathbf{d}}_1 \tag{3.17}$$

$$\mathbf{p}_{\text{wrist}} = \mathbf{p}_{\text{elbow}} + L_2\, \hat{\mathbf{d}}_2 \tag{3.18}$$

$$\mathbf{p}_{\text{tip}} = \mathbf{p}_{\text{wrist}} + L_3\, \hat{\mathbf{d}}_3 \tag{3.19}$$

and the scatterer center for each segment is the midpoint of its start and end joints. The left wing is obtained by inverting the y-component: $\hat{\mathbf{d}}_{k,\text{left}} = [\hat{d}_{k,x},\ -\hat{d}_{k,y},\ \hat{d}_{k,z}]^T$.

The effect of lead-lag on the wing planform is illustrated in Fig. 3.5, which shows the top-down wing shape for three sweep configurations.

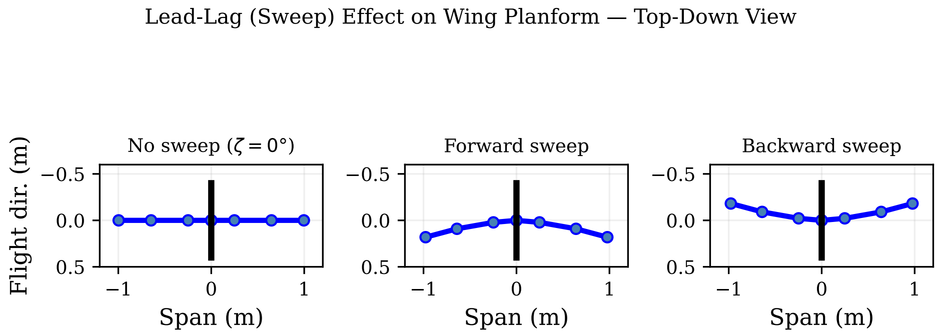


Figure 3.5: Effect of lead-lag (sweep) on the wing planform viewed from above. Backward sweep causes the distal segments to trail the proximal ones, modifying the aspect angle to the radar and tilting the micro-Doppler ridges.

3.5.3 Body Pitch Oscillation and Vertical Heave

High-fidelity flight simulations [24], [26] show that the bird body is not kinematically stationary during flapping: it oscillates in pitch (nose up/down) and undergoes vertical heave, both at the wing-beat frequency. These body oscillations are modeled as:

$$\Delta\theta_{\text{body}}(t) = \theta_{\text{pitch}}\, g_f(\tau;\, \eta_d) \tag{3.20}$$

$$\Delta z_{\text{body}}(t) = z_{\text{heave}}\, g_f(\tau;\, \eta_d) \tag{3.21}$$

where $\theta_{\text{pitch}} = 4°$ and $z_{\text{heave}} = 15\,\text{mm}$ are typical values for a medium-sized bird [26]. The pitch oscillation is added to the heading pitch angle before constructing the body rotation matrix, and the heave is added to the global z-coordinate. Figure 3.6 shows both components over two flap cycles.

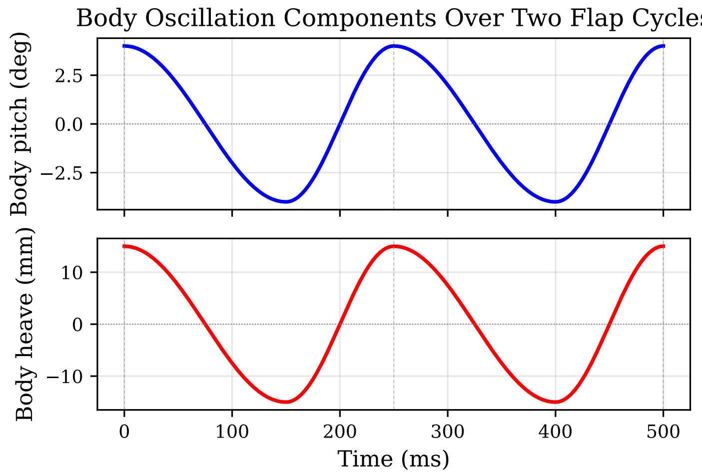


Figure 3.6: Body pitch and heave oscillation over two flap cycles. Both follow the asymmetric stroke function, with upward motion during the downstroke power phase.

The body-to-global rotation matrix is constructed from heading yaw (ψ), heading pitch (γ_0), and the body pitch oscillation ($\Delta\theta_{\text{body}}$):

$$\mathbf{R} = \mathbf{R}_z(\psi)\,\mathbf{R}_y\Big(-(\gamma_0 + \Delta\theta_{\text{body}})\Big) \tag{3.22}$$

A scatterer at body-frame position $\mathbf{p}_{\text{local}}$ is projected to global coordinates as:

$$\mathbf{p}_{\text{global}}(t) = \mathbf{R}(t)\,\mathbf{p}_{\text{local}}(t) + \mathbf{r}_{\text{bird}}(t) + [0,\ 0,\ \Delta z_{\text{body}}(t)]^T. \tag{3.23}$$

3.5.4 Stroke-Dependent Twist

The pronation/supination angle of each wing segment differs between the downstroke and upstroke [24]. For the forearm and hand segments, respectively:

$$\theta_k(t) = \begin{cases} \theta_{k,\text{down}}, & \tau < \eta_d \\ \theta_{k,\text{up}}, & \tau \geq \eta_d \end{cases} \qquad k \in \{2, 3\} \tag{3.24}$$

In the current implementation, twist affects the orientation of the scatterer's principal axis relative to the radar LOS, influencing the aspect-dependent RCS. The effect becomes physically significant when the ellipsoidal scatterers are non-circular ($a \neq b$); in the present baseline study, the scatterers are circular in cross-section ($a = b$), so the twist is stored for forward compatibility without altering the current results. Typical values from [24] are $\theta_{3,\text{down}} = 20^\circ$, $\theta_{3,\text{up}} = -30^\circ$ for the hand segment.

3.5.5 Optional Intermittent Flight

Many bird species—particularly passerines—employ an intermittent "flap-and-pause" flight strategy [21]. The bird alternates between active flapping bursts of duration T_f and passive pauses of duration T_p, during which the wings are partially folded, and the body follows a ballistic trajectory. This behavior is modeled by a gating function:

$$m_b(t) = \begin{cases} 1, & (t \bmod T_b) < T_f \\ 0, & \text{otherwise} \end{cases} \qquad T_b = T_f + T_p \tag{3.25}$$

During pauses ($m_b = 0$), the flap and sweep angles are interpolated to predetermined pause postures, and the body oscillation is suppressed. In the simulation, smooth transitions are ensured by a moving-average filter over the gating boundaries to prevent spectral artifacts. Figure 3.7 illustrates the gating waveform and gated stroke function.

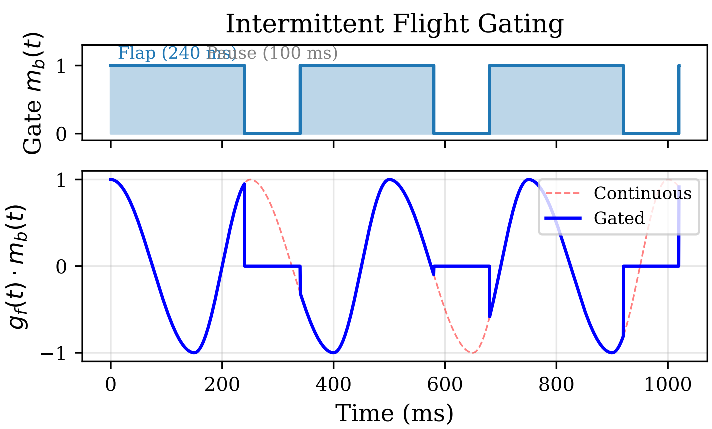


Figure 3.7: Intermittent flight gating with $T_f = 240\,\text{ms}$ flapping and $T_p = 100\,\text{ms}$ pause. During pauses, the wings fold, and the stroke function is suppressed.

3.5.6 Aspect-Dependent Ellipsoid RCS

Each scatterer is modeled as a triaxial ellipsoid with semi-axes (a, b, c), where c is along the segment's principal axis and (a, b) are perpendicular. The radar cross section is [27]:

$$\sigma(\theta) = \frac{\pi\, a^2\, b^2\, c^2}{\left(a^2 \sin^2\theta + c^2 \cos^2\theta\right)^2} \tag{3.26}$$

where θ is the aspect angle between the radar LOS and the segment's principal axis, computed via the dot product:

$$\cos\theta = \hat{\mathbf{n}}_{\text{LOS}} \cdot \hat{\mathbf{c}}_{\text{seg}}. \tag{3.27}$$

Table 3.2 summarises the RCS parameters for each scatterer.

Table 3.2: Ellipsoid RCS parameters for each scatterer. Values are simulation design choices selected to match the medium-sized-bird scatterer dimensions used in the model; the ellipsoid RCS form is taken from [27].

Scatterer	a (m)	b (m)	c (m)	Count
Body	0.10	0.10	0.40	1
Arm	0.05	0.05	0.125	2 (L+R)
Forearm	0.04	0.04	0.20	2 (L+R)
Hand	0.03	0.03	0.175	2 (L+R)
Total scatterers per bird				**7**

3.5.7 Radar Return Signal Model

The complex baseband return from a single bird is the coherent sum over all $Q = 7$ scatterers at each pulse n:

$$s[n] = \sum_{q=1}^{Q} \sqrt{\sigma_q[n]}\ \exp\left(-j\,\frac{4\pi\, r_q[n]}{\lambda}\right) \tag{3.28}$$

where $r_q[n]$ is the range from the radar to the q-th scatterer at pulse n, and $\sigma_q[n]$ is its aspect-dependent RCS. The return is accumulated into range bins of width $\Delta r = 18\,\mathrm{m}$ to form a range–time matrix.

3.5.8 Summary of Model Parameters

Table 3.3 lists the kinematic parameters used in the three-segment simulation.

Table 3.3: Three-segment bird-model parameters. Values are simulation design choices informed by the cited flapping-flight literature; they are not a fitted species model.

Parameter	Symbol	Value
Arm length	L_1	0.25 m
Forearm length	L_2	0.40 m
Hand length	L_3	0.35 m
Total half-span	$L_1 + L_2 + L_3$	1.00 m
Flap amplitudes	A_1, A_2, A_3	$45°,\ 30°,\ 25°$
Flap offsets	$\xi_{1,0},\ \xi_{2,0},\ \xi_{3,0}$	$10°,\ 25°,\ 15°$
Downstroke fraction	η_d	0.60
Wing-beat frequency	f_w	4.0 Hz
Lead-lag amplitudes	$\hat{\zeta}_1,\ \hat{\zeta}_2,\ \hat{\zeta}_3$	$5°,\ 10°,\ 15°$
Body pitch amplitude	θ_{pitch}	4°
Body heave amplitude	z_{heave}	15 mm
Intermittent (optional)	$T_f,\ T_p$	240 ms, 100 ms

3.5.9 Herring Gull Parameterization

The Herring Gull (*Larus argentatus*) provides a biological reference case for the wing-beat frequency and approximate scale used in the simulation. Table 3.4 reports the literature values and the selected parameters side by side so that departures from the reference are explicit.

Table 3.4: Herring Gull (*Larus argentatus*) morphometrics vs. simulation parameters.

Parameter	Herring Gull (literature)	Simulation
Body mass	0.8 kg to 1.2 kg [21]	—
Full wingspan	1.24 m to 1.55 m [21]	2.0 m
Body length	∼0.55 m	0.80 m
Wing-beat frequency	3.5 Hz to 4.5 Hz [21]	$\mathcal{N}(4.0,\ 0.3^2)$ Hz
Cruise speed	10 m/s to 15 m/s	$\mathcal{N}(12,\ 0.5^2)$ m/s
Intermittent flight	Observed [21]	$T_f = 240$ ms, $T_p = 100$ ms

The simulated full span of 2.0 m exceeds the cited 1.24 m to 1.55 m Herring

Gull range, and the simulated body length likewise exceeds the tabulated reference. These values were selected to increase separation of component Doppler trajectories in the simulation. The model should therefore be described as a Herring-Gull-informed parameterization, not as a morphometrically fitted or biologically validated Herring Gull model.

3.6 Reynolds-Rule Flock Trajectory Model

To examine multi-bird superposition, the three-segment wing model is incorporated into a Reynolds-rule flocking simulation [10]. Each bird is represented as an autonomous agent whose acceleration contains four algorithmic terms:

- **Separation:** Repulsive force inversely proportional to distance from neighbours within radius d_{rep}. Prevents collisions and maintains inter-bird spacing.
- **Alignment:** Velocity-matching force toward the mean heading of neighbours within radius d_{align}. Produces coordinated heading changes.
- **Cohesion:** Attraction toward the centroid of the local neighborhood. Keeps the flock together.
- **Speed regulation:** Soft clamping to a per-bird nominal speed drawn from $\mathcal{N}(v_{\text{ideal}}, \sigma_v^2)$.

A stochastic perturbation $a_{\text{rand}}\mathcal{N}(0, 1)$ is added to each acceleration component. The total acceleration is limited to $a_{\max}$, and speed is constrained to $[v_{\min}, v_{\max}]$. These are simulation rules rather than an identified model fitted to measured flock trajectories.

3.6.1 Flock Initialization

The flock is initialized in a V-formation with half-angle $\alpha_V = 30°$ and inter-bird spacing of 3 m, centered at [0, 100, 50] m. Each initial velocity is directed along the

$+x$ axis with the bird's assigned nominal speed. The Reynolds citation supports the agent-based flocking rules; the specific V angle and spacing are simulation design parameters rather than values derived from that model.

3.6.2 Flocking Parameters

Table 3.5 lists the complete set of Reynolds-rule and per-bird stochastic parameters.

Table 3.5: Reynolds-rule flocking simulation parameters.

Parameter	Symbol	Value
Number of birds	N_d	10
V-formation half-angle	α_V	30°
Inter-bird spacing	d_{space}	3 m
Formation centre	—	[0, 100, 50] m
Repulsion radius	d_{rep}	3 m
Repulsion gain	a_{rep}	2
Alignment radius	d_{align}	8 m
Alignment gain	a_{align}	0.3
Cohesion gain	a_{coh}	0.5
Random perturbation	a_{rand}	0.3
Max acceleration	$a_{\max}$	$5\,\text{m/s}^2$
Speed range	$[v_{\min}, v_{\max}]$	[8, 16] m/s
Ideal cruising speed	v_{ideal}	$\mathcal{N}(12, 0.5^2)$ m/s
Wing-beat frequency	f_w	$\mathcal{N}(4.0, 0.3^2)$ Hz
Initial flap phase	ϕ_0	$\mathcal{U}(0, 2\pi)$

Each bird is assigned an independent wing-beat frequency, nominal speed, and initial flap phase, producing controlled per-bird variability. The trajectory integrator runs at the radar PRF (2757.8 Hz), so the numerical state is updated once per simulated pulse.

3.6.3 Per-Bird Doppler Diversity

Because each bird follows a slightly different flight trajectory, the flock produces a superposition of distinct micro-Doppler signatures. The per-bird bulk Doppler is:

$$f_{\text{bulk},d}(t) = \frac{2}{\lambda} \frac{(\mathbf{r}_{\text{radar}} - \mathbf{r}_d(t)) \cdot \mathbf{v}_d(t)}{\|\mathbf{r}_{\text{radar}} - \mathbf{r}_d(t)\|}. \tag{3.29}$$

Table 3.6 lists the per-bird stochastic parameters and initial bulk Doppler offsets for all ten birds in the representative simulation run (rng(42)).

Table 3.6: Per-bird stochastic parameters and initial bulk Doppler for all ten birds in the 10-bird Herring Gull flock (rng(42)). Radar at [0, 0, 3] m; formation centre at [0, 100, 50] m. Source: `BirdModel/FlockDopplerSim/FlockDopplerSim.m` (line 37).

Bird	f_w (Hz)	v_{ideal} (m/s)	ϕ_0 (rad)	V-position	$f_{\text{bulk},0}$ (Hz)
1	3.839	11.704	0.876	Lead	0.0
2	4.260	12.922	1.836	Left-1	-18.8
3	4.293	12.908	2.302	Right-1	-19.3
4	4.101	11.938	2.866	Left-2	-34.4
5	3.701	11.445	4.933	Right-2	-34.6
6	3.843	11.660	1.255	Left-3	-49.7
7	3.611	12.007	3.231	Right-3	-55.1
8	4.275	11.970	3.722	Left-4	-67.1
9	4.053	11.669	0.292	Right-4	-72.1
10	4.227	12.153	3.817	Left-5	-84.0

The spread in wing-beat frequency ($\sim$3.6 Hz to 4.3 Hz) means that the periodic micro-Doppler ridges from different birds beat at slightly different rates, producing a richer composite spectrogram than a single bird. The initial bulk Doppler offsets span ~ -84 Hz to 0 Hz: the lead bird has zero radial velocity at $t = 0$ because it is initialized perpendicular to the radar LOS, while the trailing wingmen acquire progressively more negative bulk Doppler as their positions step further back along the $-x$ axis. These effects are collectively referred to as *swarm Doppler* *diversity* and are visible in the per-bird and composite spectrograms presented in the following section.

3.7 Simulated Micro-Doppler Results

3.7.1 Radar and Scenario Parameters

All simulations use the Furuno FAR-15x8 X-band marine radar parameters:

- Carrier frequency: $f_0 = 9.41\,\text{GHz}$ ($\lambda = 3.18\,\text{cm}$)
- Pulse repetition frequency: $\text{PRF} = 2757.8\,\text{Hz}$
- Number of pulses: $N_p = 16384$ ($\text{CPI} \approx 5.94\,\text{s}$)
- Range resolution: $\Delta r = 18\,\text{m}$
- Radar position: [0, 0, 3] m
- Flock formation centre: [0, 100, 50] m
- STFT window: Kaiser ($\beta = 10$, $M = 256$), overlap 250, $N_{\text{FFT}} = 512$

The spectrograms are displayed without bulk-motion compensation, so the micro-Doppler pattern for each bird is superimposed on its individual translational Doppler offset $f_{\text{bulk},d}$ as defined in Eq. (3.29). This preserves a measurement-consistent representation of the observed Doppler process: a coherent radar measures the total phase change from both translation and wing motion, and the per-bird bulk offsets in Table 3.6 are therefore directly visible in the displayed spectrograms.

3.7.2 Doppler Contributions by Scatterer

Figure 3.8 shows the instantaneous Doppler frequency contribution of each scatterer over one flap cycle. The body produces a nearly constant Doppler line, while the arm, forearm, and hand segments produce oscillating Doppler trajectories with increasing amplitudes. The hand segment, being the most distal, generates the widest Doppler excursion, confirming the importance of adding a third wing segment.

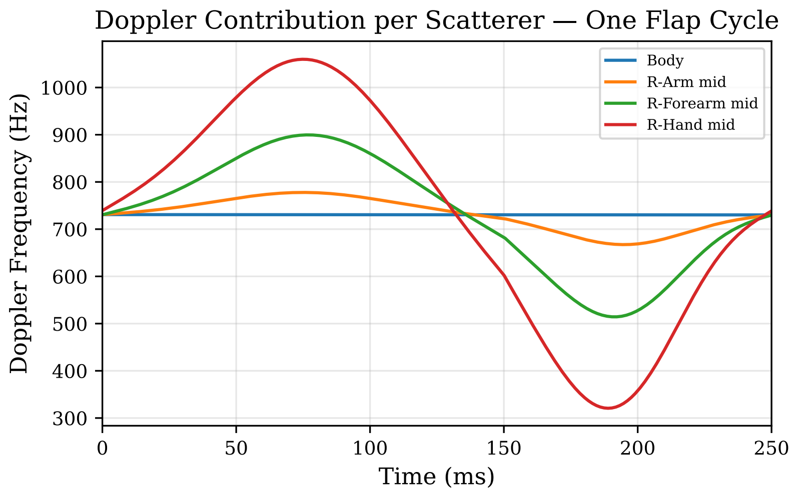


Figure 3.8: Doppler frequency contribution of each scatterer (body, arm, forearm, hand) over one flap cycle for a single bird. The hand segment produces the largest micro-Doppler extent.

3.7.3 Wing-Tip Trajectory

The three-dimensional trajectory of the elbow, wrist, and wingtip joints over one complete flap cycle is shown in Fig. 3.9. The trajectories are non-planar due to the lead-lag motion, and the downstroke–upstroke transition produces a visible asymmetry in the path geometry.

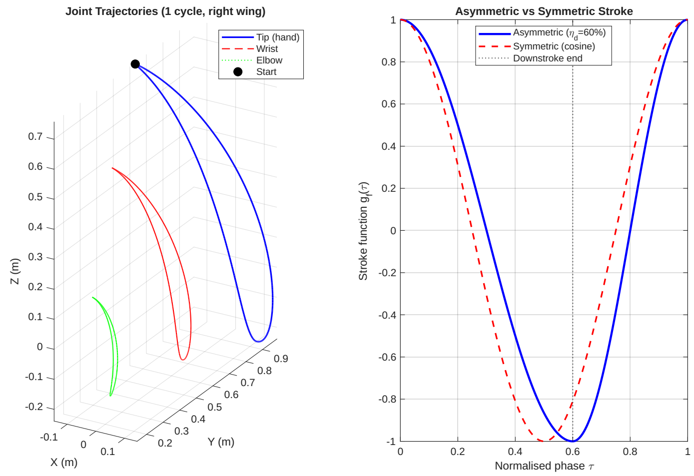


Figure 3.9: Three-dimensional trajectories of the elbow, wrist, and wingtip joints over one flap cycle. The asymmetric stroke and lead-lag produce non-planar, temporally asymmetric paths.

3.7.4 Aspect-Angle Dependence of Micro-Doppler Signatures

The micro-Doppler signature of a bird is heavily dependent on the radar viewing geometry. To quantify this, the return from bird 1 in the flock simulation is recomputed from four synthetic radar positions, each at 100 m from the formation center but at different aspect angles relative to the bird's average flight direction ($+x$). All other parameters, wing kinematics, flap frequency, and intermittent gating, remain identical, isolating the effect of observation geometry.

Because the flapping stroke is, to first order, a rotation in the body y-z plane, the instantaneous wing-tip velocity vector is dominated by its *vertical* ($\hat{z}$) component. The sweep/lead-lag oscillation ($\hat{x}$) and the body-pitch oscillation are secondary contributions. Consequently, the dominant micro-Doppler bandwidth scales with $|\hat{z} \cdot \hat{\mathbf{n}}_{\text{LOS}}|$, while the bulk Doppler scales with $|\hat{x} \cdot \hat{\mathbf{n}}_{\text{LOS}}|$ for a bird flying along $+x$. The four viewing geometries (Fig. 3.10) are:

1. **Head-on ($\theta_a \approx 0°$):** Radar directly ahead of the flight path. The bird

approaches the radar at its maximum radial velocity, producing the *largest* bulk Doppler offset. The vertical wing-tip motion is nearly perpendicular to the LOS ($\hat{z} \cdot \hat{\mathbf{n}} \approx 0$); the residual micro-Doppler modulation visible in the spectrogram arises from lead-lag sweep, body-pitch oscillation, and the small horizontal excursion of the wing-tip as the flap angle changes.

2. **Broadside ($\theta_a \approx 90°$):** Radar to the side of the flight path. Both $\hat{x} \cdot \hat{\mathbf{n}}$ and $\hat{z} \cdot \hat{\mathbf{n}}$ are small, so the bulk Doppler is near zero and the micro-Doppler bandwidth is *narrow*. A slow drift of the bulk line from slightly positive toward slightly negative Doppler is produced as the bird passes the radar's cross-range axis.
3. **Oblique ($\theta_a \approx 45°$):** An intermediate geometry with moderate bulk Doppler and a narrow-to-moderate micro-Doppler extent. This is representative of many practical surveillance scenarios.
4. **Overhead:** Radar directly above the bird. The LOS is aligned with the dominant (vertical) wing-tip velocity component, so the micro-Doppler bandwidth is *maximized* and dominates the return. The horizontal flight velocity is nearly perpendicular to the LOS, so the bulk Doppler remains small.

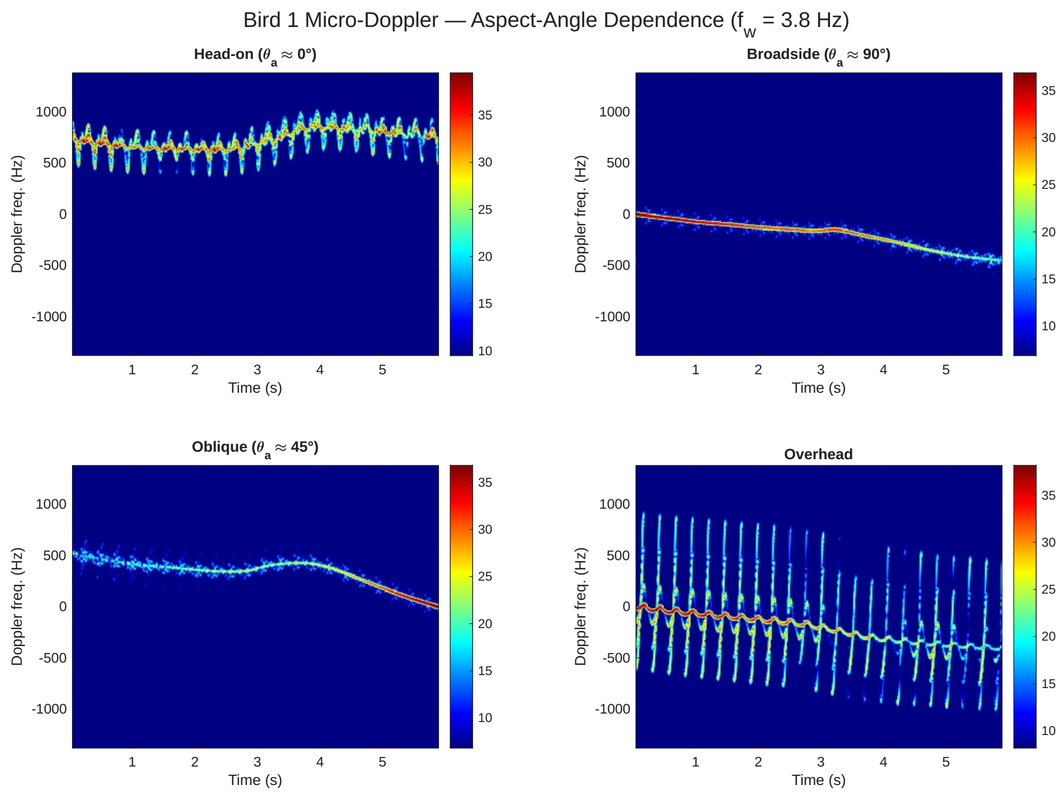


Figure 3.10: Micro-Doppler spectrograms of bird 1 viewed from four radar positions at different aspect angles. Head-on: maximum bulk Doppler, narrow micro-Doppler extent produced by lead-lag sweep and body-pitch oscillation. Broadside: near-zero bulk Doppler and narrow micro-Doppler extent (both $\hat{x}$ and $\hat{z}$ weakly project onto the LOS). Oblique: intermediate bulk Doppler and narrow-to-moderate micro-Doppler extent. Overhead: vertical wing-flap velocity projects fully onto the LOS, producing the widest micro-Doppler bandwidth ($\sim \pm 1$ kHz) with near-zero bulk Doppler. Dynamic range: 30 dB.

The key mechanism is the projection of each scatterer's velocity onto the radar LOS. For a wing segment oscillating with peak velocity v_{tip} at angle α to the LOS, the observed micro-Doppler shift is

$$f_{\text{mD}} = \frac{2\, v_{\text{tip}} \cos \alpha}{\lambda}. \tag{3.30}$$

Because the flap motion is predominantly vertical in the body frame, $\cos \alpha$ is largest when the LOS itself is vertical. The overhead geometry therefore produces the widest micro-Doppler extent, appearing in the spectrogram as periodic high-amplitude excursions reaching $\sim \pm 1\,\mathrm{kHz}$ superimposed on a near-zero bulk line. In the head-on and broadside geometries the LOS is nearly horizontal; the vertical flap velocity projects only weakly, so the micro-Doppler bandwidth is narrow. The residual modulation observed in the head-on spectrogram is driven by lead-lag sweep, body-pitch oscillation, and the horizontal excursion of the wing-tip as the flap angle varies, rather than by the flap itself. The oblique geometry is intermediate in both bulk Doppler and micro-Doppler extent.

In addition to micro-Doppler bandwidth, the aspect angle modulates each scatterer's radar cross section through Eq. (3.26). When the segment's principal axis is aligned with the LOS ($\theta \approx 0$), the ellipsoid presents its minimum cross-section; when viewed broadside ($\theta \approx 90^\circ$), the cross-section increases, making the wing segments more visible relative to the body. This RCS variation introduces an amplitude modulation that compounds the frequency modulation described above.

These geometric effects have practical implications:

- **Classification robustness:** A classifier trained on spectrograms from a single aspect angle may fail when the bird approaches from a different direction. Training and evaluation data should include multiple viewing geometries to evaluate robustness.
- **Swarm diversity:** Within a flock, different birds occupy different positions relative to the radar, each experiencing a slightly different effective aspect angle. This is one source of the per-bird spectrogram diversity visible in Fig. 3.15 and quantified in Table 3.6.
- **Time-varying geometry:** As the bird translates, the aspect angle evolves continuously, causing both the bulk Doppler and the micro-Doppler bandwidth to change within a single CPI. The broadside geometry in particular

produces a visible bulk-Doppler transition from positive to negative as the bird passes the radar's cross-range axis.

3.7.5 Single-Bird Spectrogram

Figure 3.11 shows the micro-Doppler spectrogram of bird 1 extracted from the flock simulation. This spectrogram is computed from the isolated per-bird range-bin returns (before summation across birds) using the same STFT parameters as the composite and per-bird figures, preserving consistent radar geometry ($\mathbf{r}_{\text{radar}} = [200, 0, 50]$ m) and the exported single-bird CPI length ($N_p = 8192$).

The spectrogram exhibits the following features from the three-segment model of Section 3.5:

1. **Temporal asymmetry:** Unequal dwell times for the downstroke and upstroke ridges, indicating the 60/40 asymmetric stroke function (Eq. (3.13)).
2. **Extended bandwidth:** The hand (primary) segment extends the micro-Doppler range beyond what a two-segment model produces.
3. **Ridge skew:** Lead-lag motion introduces cross-coupling between span-wise and fore-aft velocity components, producing subtle Doppler ridge asymmetry.
4. **Harmonic richness:** The asymmetric stroke and body oscillation enrich the harmonic structure relative to the symmetric-cosine Chen baseline.

Intermittent flight is disabled in this run ($m_b \equiv 1$), so burst-pause gating is not present in the figure; its effect is discussed separately in Section 3.5.5.

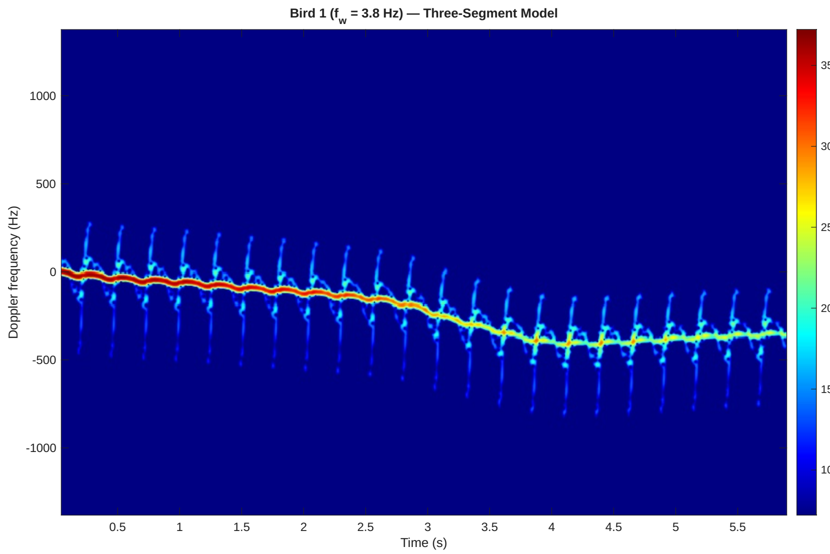


Figure 3.11: Micro-Doppler spectrogram of bird 1 from the flock simulation, using the three-segment wing model with asymmetric stroke, lead–lag motion, and body oscillation; intermittent flight is disabled for this run. Radar at [200, 0, 50] m; $N_p = 8192$; dynamic range: 30 dB.

3.7.6 Flock Formation Evolution

Figure 3.12 shows the flock at the initial, mid-CPI, and final simulation times. The Reynolds-rule dynamics transform the initialized V-formation while the cohesion term limits dispersion. Velocity vectors indicate the heading changes generated by the alignment, separation, cohesion, and speed-limiting rules.

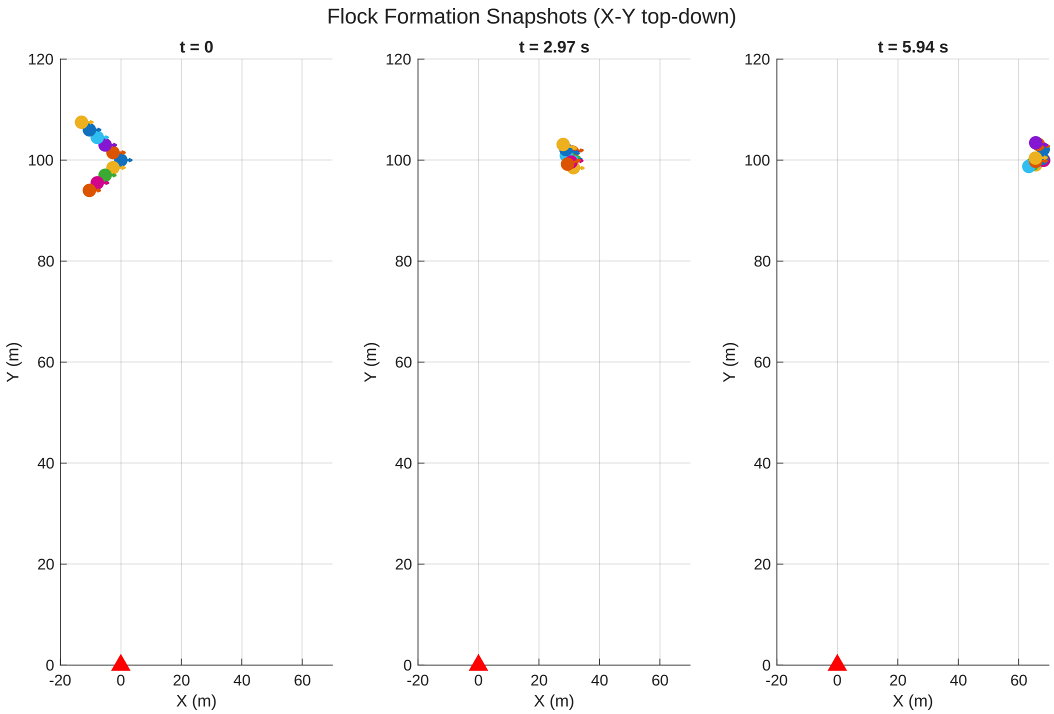


Figure 3.12: Flock formation snapshots at $t = 0$, $t \approx 2.97\,\mathrm{s}$, and $t \approx 5.94\,\mathrm{s}$. Circles mark bird positions; arrows indicate velocity direction and magnitude. The red triangle marks the radar position.

3.7.7 Per-Bird Range Traces

Figure 3.13 plots the range from the radar to each bird's body center over the full CPI. Because the birds are initialized in a V-formation with the lead bird nearest to the radar along the approximate line of sight, the range traces are separated at $t = 0$ and evolve as the flock moves. The non-parallel traces reflect the per-bird velocity differences and flocking-induced heading adjustments.

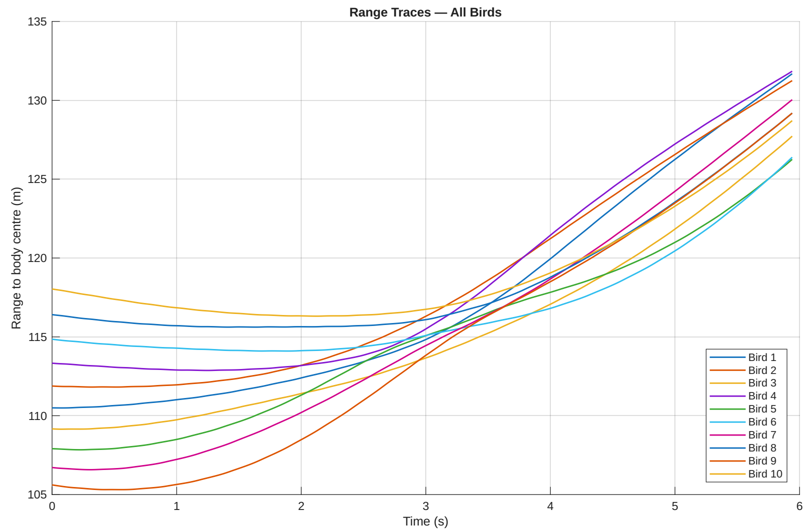


Figure 3.13: Range from radar to each bird's body center over the CPI. Different birds occupy different range cells and exhibit different range rates, adding to the Doppler diversity described in Table 3.6.

3.7.8 Composite Swarm Spectrogram

Figure 3.14 shows the STFT spectrogram of the coherent composite radar return summed across all range bins without bulk-motion compensation. The micro-Doppler patterns from the ten birds are each centered at their individual bulk Doppler offset and modulated by their respective wing-beat frequency and phase.

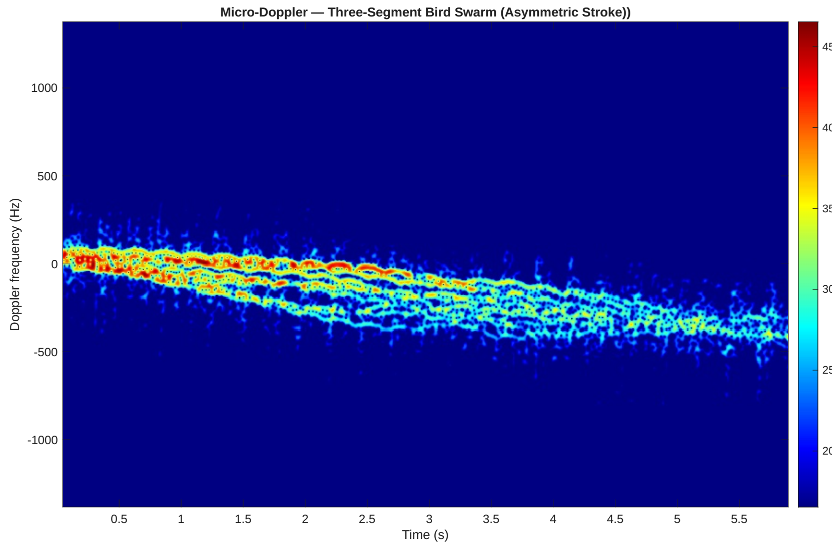


Figure 3.14: Composite swarm micro-Doppler spectrogram for the 10-bird Herring Gull flock. The superposition of individually offset, phase-diverse, and kinematically distinct returns produces a wider and more complex signature than any single bird. Dynamic range: 30 dB.

Several features distinguish the swarm spectrogram from the single-bird case:

1. **Doppler spread:** The per-bird bulk Doppler offsets (~ -84 Hz to 0 Hz initially, growing as the flock disperses) smear the body return across a wider Doppler band than a single bird would occupy.
2. **Beating:** The slightly different wing-beat frequencies produce amplitude modulation in the composite return, visible as time-varying ridge intensities.
3. **Trajectory diversity:** The independent flocking trajectories introduce time-varying phase and amplitude differences that make the composite signature broader and less regular than any one bird alone.

3.7.9 Per-Bird Spectrograms

Figure 3.15 presents the individual micro-Doppler spectrogram of each bird in the flock, computed from its isolated range-bin returns. Each bird's spectrogram displays the asymmetric-stroke micro-Doppler pattern, centered at its individual bulk Doppler offset and modulated by its own wing-beat frequency and flap phase. The wing-beat period varies between birds in accordance with the per-bird f_w values in Table 3.6.

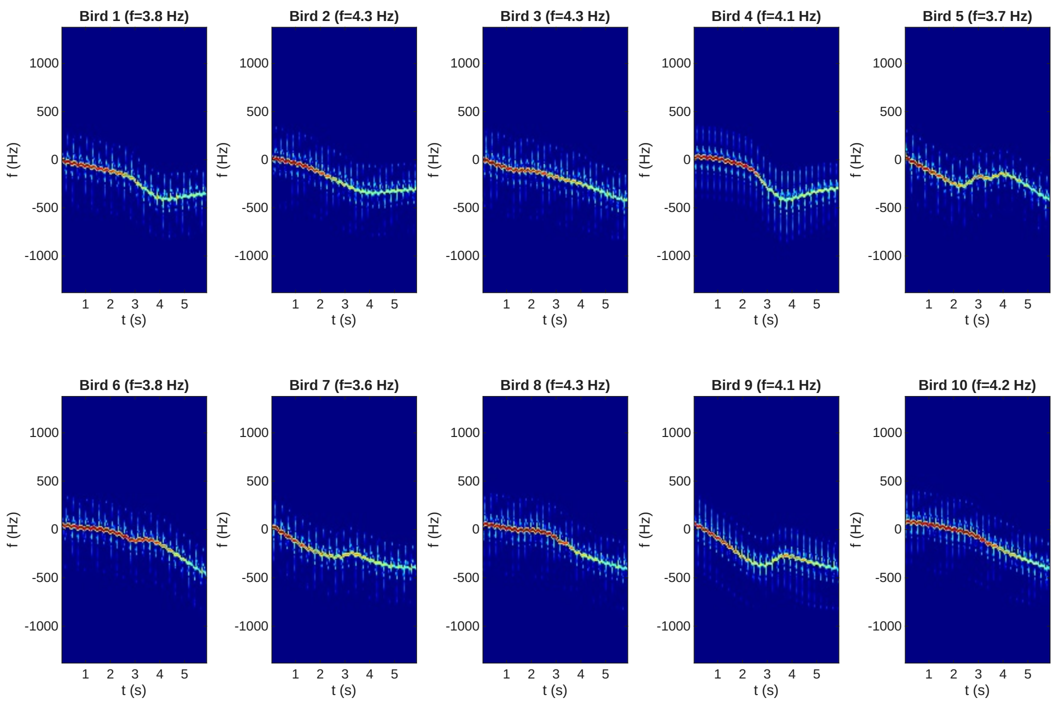


Figure 3.15: Individual micro-Doppler spectrograms for each of the 10 birds in the flock. Each bird exhibits a distinct bulk Doppler offset, wing-beat frequency, and initial phase. Dynamic range: 30 dB.

3.8 Simulated Drone Micro-Doppler Signatures

For comparison with bird signatures and for multi-class classification, drone micro-Doppler signatures are also simulated using a quadcopter model.

3.8.1 Simulation Parameters

- Quadcopter with four propellers, each with 2 blades.
- Propeller rotation rate: 3000 RPM (50 Hz).
- Propeller radius: 0.15 m.

- Body speed: hovering (0 m/s bulk Doppler) or slow movement (2 m/s).
- Body RCS: $0.1\,\mathrm{m}^2$; blade tip RCS: $0.001\,\mathrm{m}^2$ per blade.
- Initial range: 150 m.

3.8.2 Resulting Spectrogram

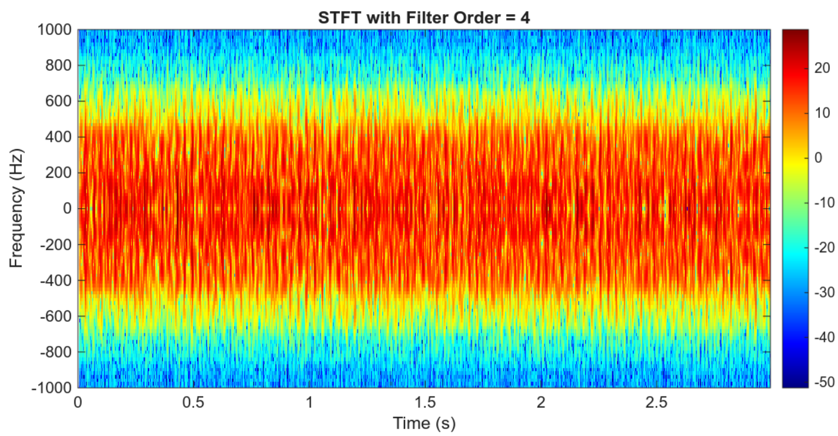


Figure 3.16: Simulated micro-Doppler spectrogram of a quadcopter drone (hovering) with propellers rotating at 3000 RPM. Radar parameters: $f_0 = 9.41$ GHz, PRF$=$ 3000 Hz.

Under the selected simulation parameters, the drone spectrogram in Fig. 3.16 differs from the bird simulation in the following ways:

- **Blade flashes:** Periodic impulse-like features extending to the maximum physical Doppler shift, determined by the blade tip speed $v_{\mathrm{tip}} = 2\pi f_{\mathrm{rot}} r_{\mathrm{blade}}$. For 3000 RPM and $r = 0.15$ m, $v_{\mathrm{tip}} \approx 47$ m/s and the pre-sampling physical value is $f_{\mu D,\max} \approx \pm 2960$ Hz. With PRF$=$ 3000 Hz, components outside ± 1500 Hz fold into the displayed slow-time Doppler interval by aliasing.
- **Flash repetition rate:** $f_{\mathrm{flash}} = N_b \times f_{\mathrm{rot}} = 2 \times 50 = 100$ Hz per propeller, yielding a characteristic high-frequency periodicity.
- **Multi-rotor superposition:** The four rotors create complex interference signatures, but the overall signature is dominated by the high-frequency

blade returns well outside the bird micro-Doppler bandwidth.

3.9 Discriminative Feature Summary

Table 3.7 summarizes the feature contrast produced by the selected bird and drone simulation parameters. The entries are simulation outcomes and literature-motivated ranges, not empirical class limits.

Table 3.7: Discriminative micro-Doppler features: bird vs. drone.

Feature	Bird (Biomechanical Micro-Doppler)	Drone (Quadcopter)
Modulation frequency	$f_w \approx 2\,\mathrm{Hz}$ to $15\,\mathrm{Hz}$	$f_{\mathrm{blade}} \approx 50\,\mathrm{Hz}$ to $200\,\mathrm{Hz}$
Max Doppler extent	$\sim$200 Hz to 600 Hz	$\sim$2000 Hz to 6000 Hz
Temporal pattern	Smooth, asymmetric sinusoidal	Impulsive blade flashes
Temporal symmetry	Asymmetric (60/40)	Symmetric (constant RPM)
Harmonic structure	Rich, non-harmonic	Harmonic comb
Body Doppler	Varies with flight speed	Near-zero (hovering) or low
Intermittent gaps	Present in passerines	Absent

These characteristics motivate candidate time–frequency features for bird–drone discrimination and the need to evaluate viewing-geometry sensitivity. The Vision Transformer experiment in Chapter 6 uses the public DIAT-μSAT benchmark rather than the simulated spectrograms from this chapter.

3.10 Chapter Summary

This chapter presented a three-segment bird micro-Doppler simulation for X-band radar. Chen's two-jointed arm formulation provides the baseline [5]; comparison with the cited flight literature motivates six kinematic and temporal modifications associated with seven baseline simplifications. The parameter set is Herring-Gull-informed but is not a morphometrically fitted species model.

The modifications are asymmetric stroke timing ($\eta_d = 0.60$), three-segment wing kinematics, lead–lag sweep, stroke-dependent twist, body pitch and heave, and optional intermittent-flight gating [21], [24]–[26]. In the simulation, these terms add component trajectories and temporal structure that are absent from the two-segment baseline. The thesis does not provide a measured-data goodness-of-fit comparison between the two models.

The three-segment model was incorporated into a Reynolds-rule simulation of ten birds. Independent trajectories, wing-beat frequencies, and initial flap phases produce the per-bird Doppler variation reported in Table 3.6. Their coherent superposition produces a broader and less regular simulated signature than the corresponding isolated single-bird return; similarity to measured multi-bird encounters is not quantified.

A systematic aspect-angle analysis showed that the micro-Doppler bandwidth depends strongly on the radar viewing geometry. Because the wing-tip velocity in this model is dominated by its vertical component, the overhead geometry aligns the dominant motion with the line of sight and produces the widest micro-Doppler bandwidth, while head-on and broadside observations yield narrow micro-Doppler extents because the vertical flap velocity projects weakly onto a near-horizontal LOS. The head-on case produces the largest bulk Doppler; the broadside and overhead cases produce near-zero bulk Doppler. This geometric dependence has clear implications for classifier robustness and for interpreting the per-bird spectrogram diversity observed within the flock.

The spectrograms are displayed without bulk-motion compensation so that translational and component-motion Doppler remain superimposed. Comparison with the selected quadcopter simulation identifies candidate feature contrasts, but the classifier in Chapter 6 is trained and evaluated on DIAT-μSAT rather than on these simulated data.

Chapter 4

Machine Learning Framework for Radar Micro-Doppler Classification

4.1 Definition and Paradigm Overview

Machine learning (ML) is a subfield of artificial intelligence focused on the development of algorithms that can learn patterns from data and make decisions or predictions without being explicitly programmed. Formally, given a set of input-output pairs $\mathcal{D} = \{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^{N}$, a machine learning model seeks to learn a function $f : \mathbb{R}^d \to \mathbb{R}^k$ that maps an input feature vector $\mathbf{x} \in \mathbb{R}^d$ to a target output $y \in \mathbb{R}^k$, such that $f(\mathbf{x}) \approx y$. The learning process involves optimizing a loss function $L(f(\mathbf{x}), y)$ over a parameter space θ.

Machine learning paradigms are commonly grouped into three categories: supervised learning, unsupervised learning, and reinforcement learning. Each category corresponds to a specific structure of the training data and a distinct learning objective.

4.2 Supervised Learning

Supervised learning encompasses a class of algorithms that utilize labeled datasets to learn a mapping $f_\theta : \mathbb{R}^d \to \mathbb{R}^k$ such that the predicted outputs $\hat{y}^{(i)} = f_\theta(\mathbf{x}^{(i)})$

closely approximate the true outputs $y^{(i)}$ for each input vector $\mathbf{x}^{(i)}$ in a dataset $\mathcal{D} = \{(\mathbf{x}^{(i)}, y^{(i)})\}_{i=1}^{N}$. The model is trained by minimizing a loss function:

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^{N} L(f_\theta(\mathbf{x}^{(i)}), y^{(i)}) \tag{4.1}$$

Loss functions typically used include:

- **Mean Squared Error (MSE)** for regression:

$$L_{\text{MSE}}(\hat{y}, y) = (\hat{y} - y)^2$$

- **Cross-Entropy Loss** for classification:

$$L_{\text{CE}}(\hat{\mathbf{p}}, y) = -\sum_{j=1}^{k} \mathbb{I}[y = j] \log(\hat{p}_j)$$

4.2.1 Regression Algorithms

Regression algorithms are used when the target variable y is continuous. These algorithms are fundamental in forecasting, function approximation, and radar-based continuous motion parameter estimation.

Linear Regression

Linear regression models the output as a weighted linear combination of input features:

$$f_\theta(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + \epsilon \tag{4.2}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is Gaussian noise. The optimal weights are derived from the closed-form normal equation:

$$\hat{\mathbf{w}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y} \tag{4.3}$$

Advantages: Fast and interpretable.

Disadvantages: Poor performance on non-linear relationships; sensitive to multicollinearity.

Bayesian Linear Regression

Unlike classical regression, Bayesian Linear Regression treats weights as probabilistic:

$$\mathbf{w} \sim \mathcal{N}(0, \Sigma_p) \tag{4.4}$$

The posterior is updated via Bayes' rule. Predictive uncertainty is represented through a Gaussian posterior predictive distribution.

Advantages: Provides confidence intervals; robust to overfitting.

Disadvantages: Computationally expensive; requires prior distribution tuning.

Hierarchical Bayesian Regression

This extends Bayesian regression by introducing group-level distributions:

$$\alpha_c, \beta_c \sim \mathcal{N}(\mu, \tau^2)$$

It models group-specific effects and is ideal for radar datasets with structured domains (e.g., multiple sensor arrays or conditions).

Advantages: Shares information across groups; improves generalization.

Disadvantages: Requires MCMC sampling; higher computational cost.

K-Nearest Neighbors Regression

This non-parametric method averages the outputs of the k nearest neighbors:

$$\hat{y} = \frac{1}{k} \sum_{i \in \mathcal{N}_k(\mathbf{x})} y^{(i)}$$

Advantages: Simple and effective locally.

Disadvantages: Poor scalability; sensitive to distance metric and dimensionality.

Random Forest Regression

Ensemble of regression trees trained on bootstrapped data. Final prediction is an average:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} f^{(t)}(\mathbf{x})$$

Advantages: Handles non-linearities; robust to noise.

Disadvantages: Less interpretable; may require tuning of tree depth and number.

Gaussian Process Regression

GP regression uses a kernel-defined covariance to define a posterior over functions:

$$\hat{y}_* = \mu_* + \Sigma_*^{1/2} \mathcal{N}(0, I) \tag{4.5}$$

$$\mu_* = K(X_*, X)(K(X, X) + \sigma^2 I)^{-1} y \tag{4.6}$$

Advantages: Provides probabilistic predictions with uncertainty bounds.

Disadvantages: $\mathcal{O}(n^3)$ complexity; kernel selection is critical.

4.2.2 Classification Algorithms

Classification problems predict discrete labels $y \in \{1, 2, \ldots, K\}$. These models are essential in radar target classification (e.g., distinguishing birds vs. drones).

Perceptron

A foundational linear classifier:

$$\mathbf{w} \leftarrow \mathbf{w} + y^{(i)}\mathbf{x}^{(i)} \quad \text{if } y^{(i)}\mathbf{w}^\top \mathbf{x}^{(i)} \leq 0$$

Advantages: Fast and intuitive.

Disadvantages: Only converges for linearly separable data; no probabilistic output.

Logistic Regression

Computes class probability using the sigmoid function:

$$P(y = 1|\mathbf{x}) = \frac{1}{1 + \exp(-\mathbf{w}^\top \mathbf{x})}$$

Extended via softmax for multi-class.

Advantages: Interpretable; probabilistic outputs.

Disadvantages: Assumes linearity in log-odds; may underfit complex data.

Naive Bayes

Assumes feature conditional independence:

$$P(y|\mathbf{x}) \propto P(y) \prod_j P(x_j|y)$$

Advantages: Scales to high dimensions; robust with small data.

Disadvantages: Unrealistic independence assumptions; poor on correlated features.

Support Vector Machines (SVM)

Maximizes margin between classes via:

$$\min \frac{1}{2}\|\mathbf{w}\|^2 \quad \text{s.t. } y_i(\mathbf{w}^\top \mathbf{x}_i + b) \geq 1$$

With kernels for non-linear decision boundaries.

Advantages: Effective on high-dimensional spaces.

Disadvantages: Sensitive to parameter selection; high memory cost for large datasets.

Decision Trees and Random Forests

Trees split data based on feature thresholds to maximize information gain. Random forests aggregate predictions from multiple trees.

Advantages: Non-parametric; interpretable.

Disadvantages: Prone to overfitting (trees); less transparent (forests).

Stochastic Gradient Descent Classifiers

SGD incrementally updates weights using mini-batches:

$$\mathbf{w} \leftarrow \mathbf{w} - \eta \nabla_\theta L(f_\theta(\mathbf{x}), y)$$

Advantages: Scalable to large datasets; online learning.

Disadvantages: Sensitive to learning rate; requires feature scaling.

Neural Networks for Classification

Neural networks model complex decision boundaries via multiple non-linear layers. The output layer typically uses softmax to produce class probabilities. These models are widely used for radar spectrogram classification when trained on micro-Doppler data [11].

$$\hat{y} = \text{softmax}(\mathbf{W}^{(L)}\phi(\ldots\phi(\mathbf{W}^{(1)}\mathbf{x})))$$

Advantages: Flexible for high-dimensional inputs (e.g., spectrograms). **Disadvantages:** Requires large datasets and tuning; difficult to interpret.

4.3 Hybrid and Semi-supervised Learning

In many applications, learning paradigms are combined. Semi-supervised learning uses a small amount of labeled data together with a larger amount of unlabeled data to improve model training. Self-supervised learning generates supervisory signals from the input data itself and has become common in deep learning workflows.

4.4 Applications in Radar and Signal Processing

Radar systems are used in defense, aviation, weather monitoring, and renewable energy operations. Deep learning has been applied to micro-Doppler classification, generation, and denoising. Radar echoes, especially those enriched with micro-Doppler effects, contain high-dimensional time-frequency structure that is amenable to modern ML methods.

4.4.1 Supervised Learning for Micro-Doppler Classification

Supervised learning has been extensively applied to classify radar returns based on micro-Doppler signatures. These signatures capture fine-grained motion features such as flapping wings (birds) or propeller blades (drones), encoded in the spectrogram of radar signals.

Yang and Zhou (2018) [11] utilized deep convolutional neural networks (CNNs) to classify human micro-Doppler signatures obtained via continuous wave radar. Their architecture consisted of convolutional layers followed by fully connected layers, trained on spectrogram images. The dataset included multiple human activities such as walking, jogging, and falling. The work reports substantial accuracy gains over classical baselines such as SVM and k-NN. The authors noted that CNNs automatically learned discriminative features without the need for handcrafted descriptors. However, they also reported limitations in generalization under varying radar geometries and occlusion scenarios.

Abdulatif et al. (2019) [12] compared deep convolutional, ensemble-tree, and classical learning approaches on range-Doppler maps for human-robot classification, with the deep convolutional network reaching the highest reported test accuracy of 99% on a single range-Doppler map (gradient-boosting ensemble: 97.85%). Despite this result, the deep approach required computational resources unsuitable for real-time edge deployment, highlighting a trade-off between performance and operational feasibility.

4.4.2 Unsupervised Learning and Clustering in Anomaly Detection

Unsupervised learning methods are employed to discover unknown patterns or anomalies in radar returns. A classical use case is clutter removal or distinguishing unexpected scatterers in surveillance radar.

An early Doppler-radar bird-classification study [28] investigated how radar return characteristics can be used to separate avian targets. While not machine learning in the modern sense, this work predates and is conceptually adjacent to later clustering approaches such as DBSCAN and k-means for segmenting radar targets based on probabilistic descriptors.

Modern approaches extend this idea through autoencoders and variational autoencoders (VAEs) trained on spectrogram data to learn latent representations. Anomalous instances, for example drones appearing in bird migration scans, manifest as reconstruction errors. Although these methods are data-efficient, they remain limited by interpretability concerns and require extensive validation datasets to control false positives [29].

4.4.3 Generative Models and Data Augmentation

To address the paucity of labeled radar data, especially for rare classes like drones or insects, generative models have been applied to augment training sets.

Wang et al. (2022) [29] designed a layer-reduced Deep Convolutional Generative Adversarial Network (DCGAN) to generate synthetic human micro-Doppler spectrograms. Their generator and discriminator networks were streamlined to reduce parameter count while maintaining spectral fidelity. Training was performed on real radar captures, and synthetic samples were reported to improve classifier performance under limited-data conditions. A key limitation was the difficulty in maintaining class separability for closely related motions (e.g., walking vs. jogging).

Tang et al. (2023) [30] introduced FMNet, a feature-wise mapping network designed to enhance noisy micro-Doppler spectrograms. Unlike GANs, FMNet performs deterministic denoising by learning a latent mapping between noisy and clean spectrograms. This method was evaluated using both LOS and through-the-wall (TTW) radar data. FMNet was reported to improve downstream classification accuracy when used as a preprocessing step. The model generalized across simulated and real spectrograms, though its performance degraded for extreme signal distortions or out-of-distribution motions.

4.4.4 Challenges and Limitations

Despite these results, ML models for radar still face several challenges:

- **Domain shift**: Models trained on simulated or laboratory data often fail in outdoor deployments due to differences in noise, multipath, and occlusion.
- **Data scarcity**: High-quality labeled radar datasets are scarce, making generalization difficult.
- **Real-time constraints**: Many deep models are too computationally intensive for onboard radar systems, necessitating model pruning or FPGA acceleration.
- **Interpretability**: Neural models provide limited insight into why a target was classified a certain way, which is an important limitation in security-sensitive contexts.

Chapter 5

Coherent-on-Receive Pipeline for X-band Non-Coherent Marine Radar

5.1 Introduction

Commercial-off-the-shelf marine navigation radars provide a comparatively inexpensive sensing platform, but a magnetron transmitter does not maintain the pulse-to-pulse phase reference assumed by coherent Doppler processing [1], [2]. For the Furuno FAR-15x8 used in this thesis, the sampled intermediate-frequency (IF) waveform for pulse m contains an uncontrolled pulse phase ϕ_m and a pulse-dependent IF $f_{\mathrm{IF},m}$. If these terms are not estimated, the slow-time phase history contains transmitter and receiver variation in addition to target motion.

This chapter documents the maintained coherent-on-receive pipeline used to generate the representative Furuno results. It focuses on the Galveston binary collections and excludes earlier exploratory ASCII captures and autofocus-style correction trials that are not part of the current CPI-based implementation. The processing is organized into a baseline COR engine and a research extension. The baseline performs the corrections applied to the primary results; the extension examines additional clutter-aware reference estimation and inter-CPI phase tracking. Several parameters and range-selection steps remain scene dependent.

The baseline engine performs adaptive IF tracking, conditional residual IF-offset correction (denoted carrier-frequency-offset, or CFO, correction in the implementation), CPI-wise leakage-reference extraction, pass-through timing with alignment diagnostics, matched filtering, and range-tracked micro-Doppler analysis. The extension adds clutter-aware leakage estimation, CPI-to-CPI phase tracking, integrated range–Doppler products, and side-by-side corrected and baseline time–frequency products. Together, these components form a software-defined COR framework for the pulse-indexed analytic IF model defined below.

5.2 Signal Model and Acquisition Context

Let $x_m[n]$ denote the complex analytic representation of the sampled IF signal at fast-time index n within pulse m. A pulse-indexed model that retains the propagation delay is

$$x_m[n] = \sum_{q=0}^{Q-1} \alpha_{q,m}\, p(nT_s - \tau_{q,m}) \exp\{j\,[2\pi f_{\mathrm{IF},m} nT_s + \phi_m - 2\pi f_{c,m}\tau_{q,m}]\} + w_m[n], \tag{5.1}$$

where T_s is the ADC sampling interval, $p(t)$ is the transmitted pulse envelope, $\alpha_{q,m}$ is a complex scattering or coupling coefficient, $\tau_{q,m}$ is the path delay, $f_{c,m}$ is the pulse carrier frequency, $f_{\mathrm{IF},m} = f_{c,m} - f_{\mathrm{LO}}$, and $w_m[n]$ is additive receiver noise. For a monostatic target, $\tau_{q,m} = 2R_q(m)/c$. One index represents the direct leakage path, with effective coupling delay $\tau_{\ell,m}$. The leakage and target terms share the pulse phase ϕ_m, so the leakage supplies the pulse-specific reference required for COR [7], [8]. The equation is evaluated at fast-time sample n; a target contributes to that sample when its delayed pulse envelope $p(nT_s - \tau_{q,m})$ has support there. Thus, the target delay determines which fast-time samples contain the target return.

After digital downconversion with a pulse-dependent estimate $\hat{f}_{\mathrm{IF},m}$, the baseband signal becomes

$$\begin{aligned}\tilde{x}_m[n] &= x_m[n]e^{-j2\pi\hat{f}_{\mathrm{IF},m}nT_s} \\ &= \sum_{q=0}^{Q-1} \alpha_{q,m}\, p(nT_s - \tau_{q,m}) \exp\{j\,[2\pi\Delta f_m nT_s + \phi_m - 2\pi f_{c,m}\tau_{q,m}]\} + \tilde{w}_m[n],\end{aligned} \tag{5.2}$$

with residual IF-frequency error $\Delta f_m = f_{\mathrm{IF},m} - \hat{f}_{\mathrm{IF},m}$. This error produces a fast-time phase slope and, when evaluated at different delays, a range-dependent phase error. IF estimation and residual-offset correction therefore precede the leakage phase rotation.

After residual IF correction, an ideal leakage phasor has phase

$$\angle r_{\ell,m} = \phi_m - 2\pi f_{c,m}\tau_{\ell,m}. \tag{5.3}$$

Rotation by $e^{-j\angle r_{\ell,m}}$ removes the pulse-common phase origin while retaining the target phase relative to the leakage path:

$$y_m[n] \propto \sum_{q=0}^{Q-1} \alpha_{q,m}\, p(nT_s - \tau_{q,m})\, e^{-j2\pi f_{c,m}(\tau_{q,m}-\tau_{\ell,m})}. \tag{5.4}$$

Thus, COR does not remove the propagation-phase change associated with target range and radial motion; it expresses that phase relative to the sampled leakage reference. Matched filtering and slow-time Doppler processing are applied after the IF and pulse-phase corrections.

Table 5.1 lists the stable system parameters used by the current scripts.

Table 5.1: Core Furuno COR processing parameters used by the maintained MATLAB implementation.

Quantity	Value
Carrier frequency f_c	9.41 GHz
Wavelength $\lambda = c/f_c$	0.0319 m
ADC sample period T_s	2.5 ns
ADC sample rate f_s	400 MHz
Nominal IF	approximately 60 MHz
Pulse duration T_p	0.12 µs
Pulse samples $N_p = \mathrm{round}(T_p/T_s)$	48
Nominal range resolution $cT_p/2$	approximately 18 m
Sample-to-range conversion $cT_s/2$	0.375 m/sample
PRF	2757.8 Hz
CPI size M	256 pulses
CPI duration $T_{\mathrm{CPI}} = M/\mathrm{PRF}$	approximately 0.0928 s
Unambiguous velocity $\pm\lambda\mathrm{PRF}/4$	approximately 21.96 m/s

5.3 Implemented Baseline COR Pipeline

5.3.1 Adaptive IF Compensation and Baseband Conversion

The current baseline script begins by estimating a coarse IF from the first pulse spectrum, restricting the search to the physically relevant 50 MHz to 70 MHz band. A quadratic interpolation around the spectral peak refines the initial estimate, after which a zero-phase bandpass filter is applied to all pulses. The code then performs a second, pulse-dependent IF estimate by tracking the dominant positive-frequency peak of the filtered spectrum and smoothing the resulting sequence with a moving median of length 31. This yields the pulse-wise downconversion frequency $\hat{f}_{\mathrm{IF},m}$ in (5.2).

Figure 5.1 shows the type of spectrum used in this step. The raw IF spectrum has energy near the positive and negative IF bands, whereas the downconverted spectrum is centered near zero frequency; the peak location provides the frequency estimate used for baseband conversion.

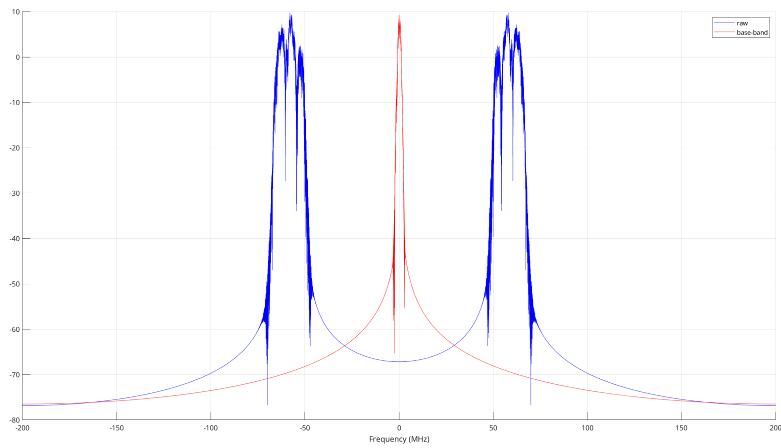


Figure 5.1: Example spectrum used for adaptive IF estimation. The raw sampled IF is shown before downconversion, and the baseband spectrum is centered near zero frequency after downconversion.

Two further implementation details matter. First, the script retains a second matched-filter-assisted IF estimate so that the leakage-reference path and the data path remain centered on the same baseband. Second, the residual IF-frequency offset is corrected only when the measured mean offset exceeds a configured threshold. In the maintained version, both the full-dwell and intra-CPI corrections are gated at 50 kHz. This guard reduces the risk that a near-DC search interprets clutter energy as receiver-frequency error.

5.3.2 Reference Peak Detection and Leakage Window Construction

The maintained pipeline no longer assumes a fixed hard-coded leakage bin. Instead, it computes the average post-trigger envelope across pulses, suppresses the pre-trigger samples, and selects the strongest persistent early-time return as the candidate leakage peak. Because COR requires a pulse-common transmit reference, the selected early-time component is treated as a leakage-path estimate rather than as a generic target echo. If the detected peak index is r_ℓ, the leakage window is defined using a half-pulse-width scale

$$\mathcal{L} = \{r_\ell - \alpha W, \ldots, r_\ell\}, \qquad W = \left\lceil \frac{N_p}{2} \right\rceil, \tag{5.5}$$

clipped to valid fast-time indices, where α is the script-configured pre-peak span. The maintained `adaptive_csvd.m` baseline uses $\alpha = 4$; related template-estimation scripts use the same half-pulse-width principle with scene-specific margins. This choice captures the leakage support together with a safety margin while keeping the reference block short enough for robust SVD-based reference estimation.

Figure 5.2 highlights the fast-time leakage region selected by the construction in (5.5).

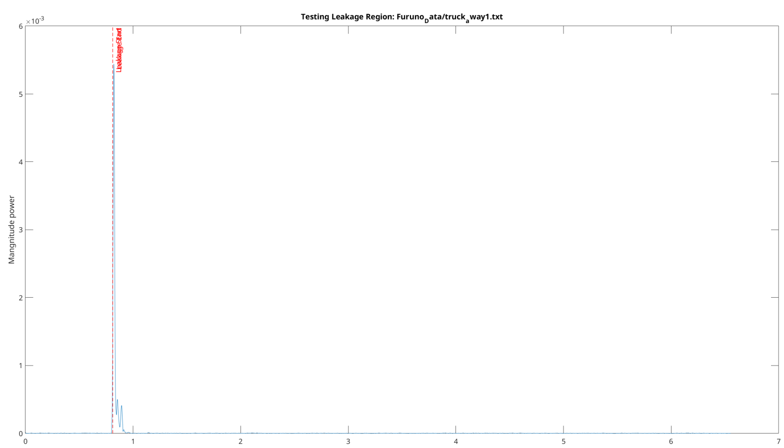


Figure 5.2: Example of the leakage-window construction. The selected leakage region is highlighted around the persistent early fast-time return used as the coherent-on-receive reference.

5.3.3 CPI-Wise Phase Restoration

The central processing unit is the CPI. For CPI k, let

$$\mathbf{L}_k \in \mathbb{C}^{M \times N_\ell} \tag{5.6}$$

denote the leakage block extracted from the CPI, where $N_\ell = |\mathcal{L}|$. The baseline script executes the following steps for each CPI.

1. Per-CPI residual CFO correction. The carrier-frequency offset (CFO) is the residual frequency difference between the pulse's actual IF and the estimated downconversion frequency. It appears as a displacement of the residual leakage spectrum away from DC and produces a fast-time phase ramp, so it must be separated from target Doppler. The CPI data are FFTed along fast time, the residual spectral peak near DC is refined by parabolic interpolation, and the slow-time sequence of CFO estimates is median- and Savitzky-Golay-smoothed. If the mean residual exceeds the 50 kHz threshold, a pulse-dependent phase ramp $e^{-j2\pi \hat{f}_{\mathrm{CFO},m} n T_s}$ is applied to the full CPI; otherwise, no CFO correction is applied.

2. Rank-1 leakage template extraction. The singular value decomposition (SVD) factors a matrix into orthonormal left and right singular vectors with non-negative singular values ordered from largest to smallest. Retaining only the largest singular value and its associated vectors gives the best rank-1 approximation in the least-squares/Frobenius-norm sense. The leakage block is therefore approximated by its dominant singular component,

$$\mathbf{L}_k \approx \sigma_{1,k}\, \mathbf{u}_{1,k} \mathbf{v}_{1,k}^H, \tag{5.7}$$

where the right singular vector $\mathbf{v}_{1,k}$ represents the within-pulse reference waveform and the left singular vector $\mathbf{u}_{1,k}$ represents its pulse-to-pulse complex coefficient. This is a practical rank-1 estimate, not a claim that the measured leakage block is exactly rank one. A matched-filter kernel is formed from the first-CPI reference and reused in the maintained implementation.

3. Fast-time alignment diagnostics. The maintained baseline keeps the phase-only-correlation (POC) alignment logic as a diagnostic and pass-through safeguard rather than applying sub-sample shifts in the reported products. Earlier tests showed that enabling leakage-window POC shifts could inject vertical stripe artifacts into the range-Doppler map when the adaptive IF tracker had already absorbed the residual range walk. The retained diagnostic forms the normalized cross-power spectrum between each pulse's leakage segment and the current CPI reference shape, locates the correlation peak, and uses local interpolation to estimate the fractional shift $\hat{\tau}_m$ in samples. If alignment is re-enabled for future scenes, it is accepted only when the shift jitter satisfies

$$\mathrm{std}(\hat{\tau}_m) \leq 1 \text{ sample}. \tag{5.8}$$

Otherwise the raw CPI is passed through unchanged. In the maintained runs used for this thesis, timing is passed through unchanged and the POC result is treated as an alignment sanity check rather than as an applied correction.

4. Pulse-wise leakage phase removal. After reference extraction and alignment diagnostics, the pulse-specific transmit phase is estimated from the average leakage phasor,

$$\hat{\phi}_m = \angle\left(\frac{1}{N_\ell}\sum_{n\in\mathcal{L}} \bar{x}_m[n]\right), \tag{5.9}$$

where $\bar{x}_m[n]$ denotes the aligned CPI data. The coherent-on-receive pulse is then

$$x_m^{(\mathrm{cor})}[n] = \bar{x}_m[n] e^{-j\hat{\phi}_m}. \tag{5.10}$$

This step explicitly removes the common magnetron phase term while preserving the target-dependent phase evolution needed for Doppler processing.

5.3.4 Matched Filtering, Range-Doppler Products, and Range Tracking

The matched filter is derived from the conjugate-reversed leakage template with a Hamming taper,

$$h[n] = w[n]\, v_1^*[N_\ell - 1 - n], \tag{5.11}$$

where $\mathbf{v}_1$ denotes the right singular vector from the first-CPI rank-1 leakage template and contains the N_ℓ fast-time reference samples. This kernel is applied to every corrected pulse. Optional MTI is implemented as mean subtraction across slow time within each CPI; this is deliberately left switchable because zero-Doppler removal can suppress stationary components that are still useful for scene understanding.

For CPI k, the matched-filtered range-Doppler map is formed as

$$R_k(r, f_d) = \sum_{m=0}^{M-1} z_k[m, r]\, w_m\, e^{-j2\pi f_d m/\mathrm{PRF}}, \tag{5.12}$$

where $z_k[m, r]$ is the matched-filter output and w_m is the slow-time Hamming window. The script stores the full CPI cube R_k for later animation and higher-level integration.

Micro-Doppler extraction is range-tracked rather than range-static. An RTI map,

$$P_k(r) = \frac{1}{M} \sum_{m=0}^{M-1} |z_k[m, r]|^2, \tag{5.13}$$

is built from the matched-filtered data. After the user selects an initial search window, the dominant range bin is tracked CPI by CPI inside an adaptive search margin of approximately 50 m, and a symmetric gate of about ±2 range-resolution cells is placed around the tracked peak. The gated slow-time signal is then analyzed with three Kaiser-window STFTs (fine, adaptive, and coarse), with 95% overlap and four-times zero-padding. Optional synchrosqueezed transforms are produced when a Fourier synchrosqueezed transform implementation is available in the signal-processing environment.

5.3.5 Diagnostic Metrics

The maintained pipeline still computes coherence diagnostics, but it does so in a way that is local to the selected CPI and scene rather than as a single thesis-wide benchmark. The key quantities are the pulse-to-pulse phase stability of a reference leakage sample,

$$\sigma_\phi = \mathrm{std}\Big(\Delta\, \mathrm{unwrap}\{\angle x_m^{(\mathrm{cor})}[n_{\mathrm{ref}}]\}\Big), \tag{5.14}$$

and the corresponding coherence factor,

$$\mathrm{CF} = \left| \frac{1}{M} \sum_{m=1}^{M} e^{j\angle x_m^{(\mathrm{cor})}[n_{\mathrm{ref}}]} \right|. \tag{5.15}$$

The advanced extension also reports inter-CPI coherence-factor comparisons and spectrogram contrast metrics, discussed in Section 5.4. Because the workflow is interactive and scene-dependent, these metrics are best interpreted together with the representative figures rather than through a single aggregate table. These metrics are computed as diagnostics for comparing phase-reference choices; they are not used as automatic selection criteria for the representative products in Section 5.5.

5.4 Advanced Extensions Beyond the Baseline COR Engine

5.4.1 Clutter-Aware Leakage Extraction

The extension begins from the pulse-phase-corrected CPI outputs of the baseline engine and estimates an alternative reference for leakage windows contaminated by sea clutter. For CPI k, the leakage matrix is modeled as

$$\mathbf{L}_k = \sigma_{\mathrm{TX}}\, \mathbf{u}_{\mathrm{TX}}\mathbf{v}_{\mathrm{TX}}^H + \mathbf{C}_{\mathrm{sea}} + \mathbf{N}, \tag{5.16}$$

where the first term is the desired rank-1 leakage component, $\mathbf{C}_{\mathrm{sea}}$ is structured sea clutter, and $\mathbf{N}$ is residual receiver noise.

The first refinement is slow-time Doppler pre-filtering around DC. For X-band sea clutter, the first-order Bragg frequency is

$$f_B = \frac{1}{2\pi}\sqrt{g\,\frac{4\pi}{\lambda}}, \tag{5.17}$$

where g is the gravitational acceleration and λ is the radar wavelength. This places the dominant gravity-wave clutter away from the exactly zero-Doppler leakage component. The implementation applies a Gaussian weight

$$W(f) = \exp\left(-\frac{f^2}{2\sigma_f^2}\right) \tag{5.18}$$

with $\sigma_f = 3\,\mathrm{Hz}$ to the slow-time Doppler spectrum of each leakage-bin column. This is not claimed as a universal optimum, but it is physically justified because the leakage is centered at DC while the clutter energy is concentrated near $\pm f_B$ [31].

The second refinement is singular-value shrinkage. If s_i are the singular values of the Doppler-filtered leakage block, with $m = \min(M, N_\ell)$, $n = \max(M, N_\ell)$, and $\beta = m/n$, the script uses the Gavish–Donoho hard-threshold rule

$$\tau^\star = \hat{\sigma}_e \sqrt{n}\, \omega(\beta), \tag{5.19}$$

where

$$\omega(\beta) = \sqrt{2(\beta+1) + \frac{8\beta}{\beta + 1 + \sqrt{\beta^2 + 14\beta + 1}}}. \tag{5.20}$$

The noise estimate $\hat{\sigma}_e$ is formed from the tail singular values. This step is mathematically better grounded than the earlier ad hoc thresholding rule, but it should still be interpreted carefully: the exact Gavish–Donoho optimality result assumes asymptotically large matrices with approximately i.i.d. Gaussian contamination, whereas the leakage block here has already been Doppler filtered and is therefore only approximately within that regime [32].

The third refinement is coherence-weighted phase extraction. For leakage sample n, the lag-1 slow-time coherence is defined as

$$\gamma(n) = \frac{\left|\sum_{m=1}^{M-1} x(m,n)x^*(m+1,n)\right|}{(M-1)\,\overline{P}(n)}, \qquad \overline{P}(n) = \frac{1}{M}\sum_{m=1}^{M} |x(m,n)|^2. \tag{5.21}$$

The resulting CPI phase reference is

$$\hat{\phi}_k(m) = \angle\left(\sum_{n\in\mathcal{L}} \gamma(n)x(m,n)\right). \tag{5.22}$$

This weighting favors leakage-window samples with larger lag-1 slow-time correlation. It is an engineering estimator whose behavior depends on the scene and is not presented as a globally optimal phase estimator.

5.4.2 Coherent CPI-to-CPI Integration

The baseline pipeline restores coherence within each CPI. The extension script adds coherence across CPIs. Let $R_k(f_d, r)$ denote the matched-filtered RDM for CPI k. If the residual CPI phase error is Φ_k, coherent accumulation takes the form

$$R_{\text{coh}}(f_d, r) = \sum_{k=0}^{K-1} R_k(f_d, r) e^{-j\hat{\Phi}_k}, \tag{5.23}$$

where K is the number of CPIs and $\hat{\Phi}_k$ is the chosen inter-CPI phase estimate.

The script evaluates three phase-reference estimators. The first is leakage-based tracking,

$$\phi_{\text{leak},k} = \angle(\text{mean}\{x_k(m, n) : n \in \mathcal{L}\}), \qquad \Delta\Phi_k = \phi_{\text{leak},k} - \phi_{\text{leak},1}, \tag{5.24}$$

which continues the COR principle by using the same leakage process as the pulse-wise phase reference [7]. The second uses the dominant zero-Doppler scatterer in the CPI RDM as a stationary-scene phase reference. The third is a cross-CPI correlation-phase estimator based on the Frobenius inner product,

$$\widehat{\Delta\Phi}_{k,k-1} = \angle(\langle R_k, R_{k-1}\rangle_F) = \angle\left(\sum_{f_d, r} R_k(f_d, r) R^*_{k-1}(f_d, r)\right). \tag{5.25}$$

This third method averages information over multiple range–Doppler cells, but changes in scene content can bias the estimate. It can also accumulate drift because phase differences are integrated sequentially rather than referenced to a fixed physical path.

If the residual phase error after tracking is ε_k, the coherent integration gain is governed by the coherence factor

$$\mathrm{CF}_{\mathrm{CPI}} = \left| \frac{1}{K} \sum_{k=0}^{K-1} e^{j\varepsilon_k} \right|, \tag{5.26}$$

so that

$$\mathrm{SNR}_{\mathrm{coh}} = K\, \mathrm{CF}_{\mathrm{CPI}}^2\, \mathrm{SNR}_{\mathrm{single}}. \tag{5.27}$$

Under the stated equal-amplitude, independent-noise model, $\mathrm{CF}_{\mathrm{CPI}} = 1$ gives an ideal coherent SNR gain of $10 \log_{10} K$ dB [2]. No universal dB rule is assumed here for post-detection power integration because its detection benefit depends on the signal and noise statistics. The extension script also includes a hybrid strategy that coherently sums within groups of CPIs and then adds group powers, for cases in which phase tracking is reliable only over shorter dwell segments.

5.4.3 Extended Micro-Doppler Analysis

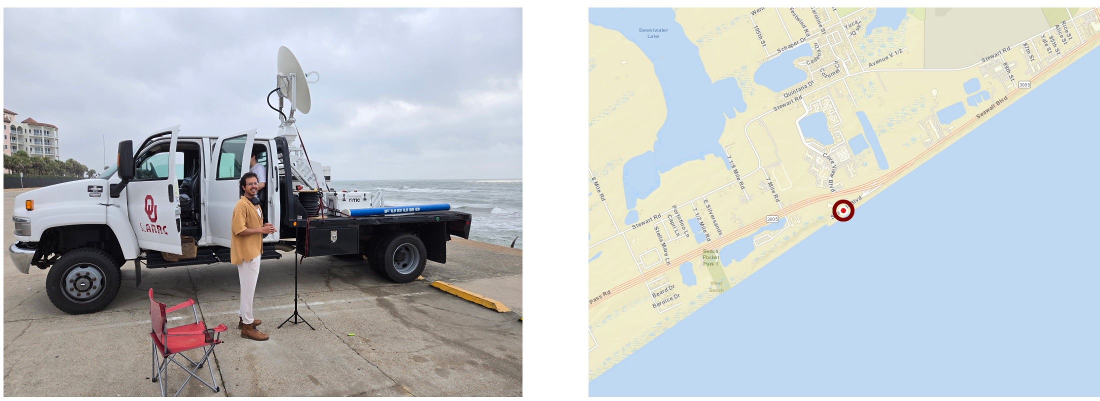
After an inter-CPI phase method is selected, the extension applies the cumulative phase correction to the matched-filtered slow-time signal to form a longer corrected dwell. Candidate targets are detected from the integrated RDM, range-tracked CPI by CPI, and analyzed with the multi-resolution STFT family used by the baseline script. Additional diagnostic products include low-rank approximations of the time–frequency power matrix, cadence–velocity diagrams (CVDs), and descriptors such as mean bandwidth, peak Doppler, Doppler asymmetry, spectral entropy, and modulation frequency estimated from autocorrelation or a CVD peak.

The extension provides side-by-side inter-CPI-corrected and baseline products. These comparisons show how the selected estimator changes a particular processed scene; they do not establish a universal numerical gain.

5.5 Representative Results

The current Furuno workflow is interactive: targets are selected from range-time-intensity (RTI) displays, range gates are tracked, and products are generated on a trial-by-trial basis. The figures below are therefore representative outputs of the maintained pipeline rather than entries of a single automated benchmark.

The field collections were acquired in a representative Galveston Seawall configuration. Figure 5.3 documents the Furuno radar platform and the geographic location of the field deployment so that the range and target descriptions below have a physical reference.



(a) Galveston shoreline collection with the Furuno radar platform.

(b) Galveston Seawall deployment location.

Figure 5.3: Furuno field setup and geographic deployment location associated with the radar collections analyzed in this chapter.

Figure 5.4 illustrates the baseline range-tracking workflow for a bird observation pointed north. The figure contains a range-tracked Kaiser STFT and a corresponding Fourier synchrosqueezed-transform (FSST) view over the same slow-time interval. The key point is that the CPI-wise range gate follows the target over time before these time-frequency representations are formed.

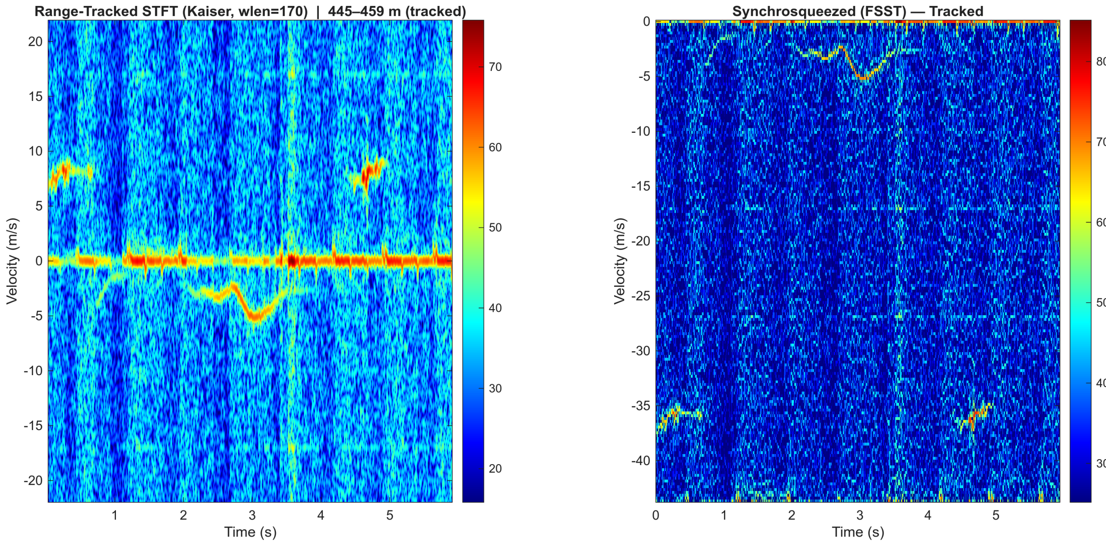


Figure 5.4: Representative bird result from the baseline COR pipeline for a northward-pointing observation collection. The two panels show a range-tracked Kaiser STFT and a Fourier synchrosqueezed-transform view over approximately 445 m to 459 m.

Figure 5.5 shows a single adaptive velocity spectrogram for a second northward-pointing bird scene. The display spans the saved range interval and shows the persistent near-zero-velocity component together with time-varying energy at later times; it is not a multi-panel corrected-versus-baseline comparison.

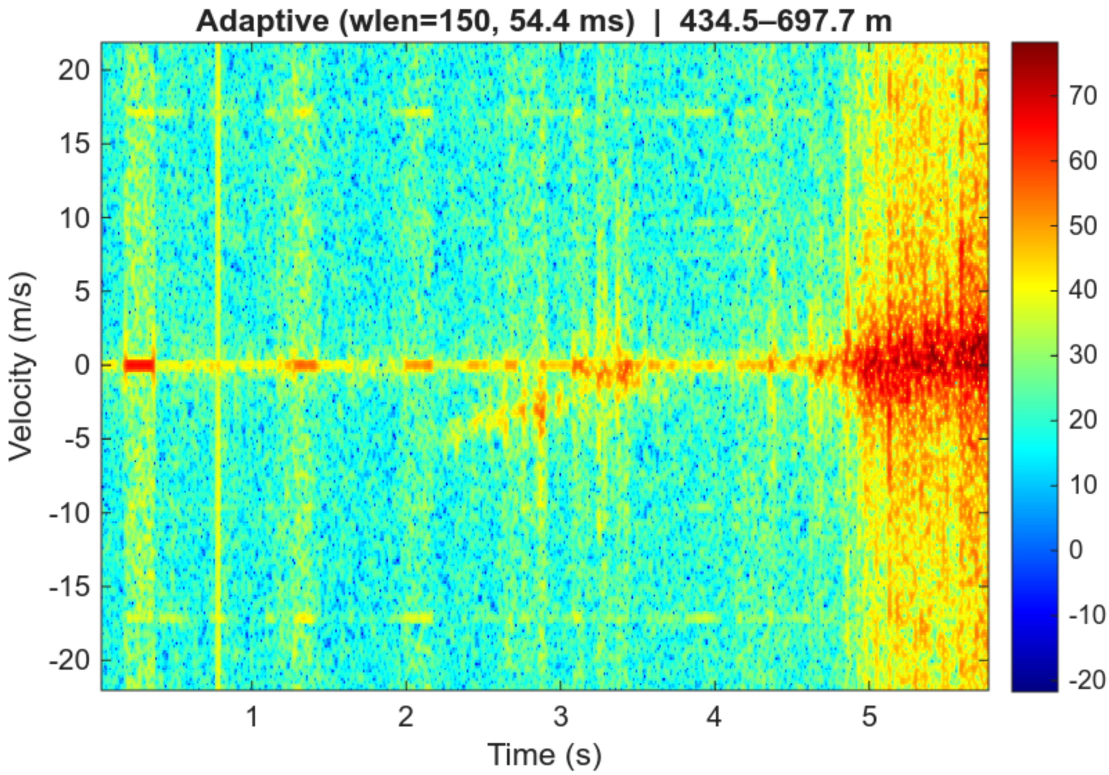


Figure 5.5: Representative adaptive velocity spectrogram for a second northward-pointing bird observation. The single panel spans the saved range interval of approximately 434.5 m to 697.7 m.

Figure 5.6 presents a single range-tracked Doppler-frequency spectrogram for a drone example collected while looking along the Galveston seawall. This example documents application of the same processing sequence to a rotary-wing target; it is not a classification experiment.

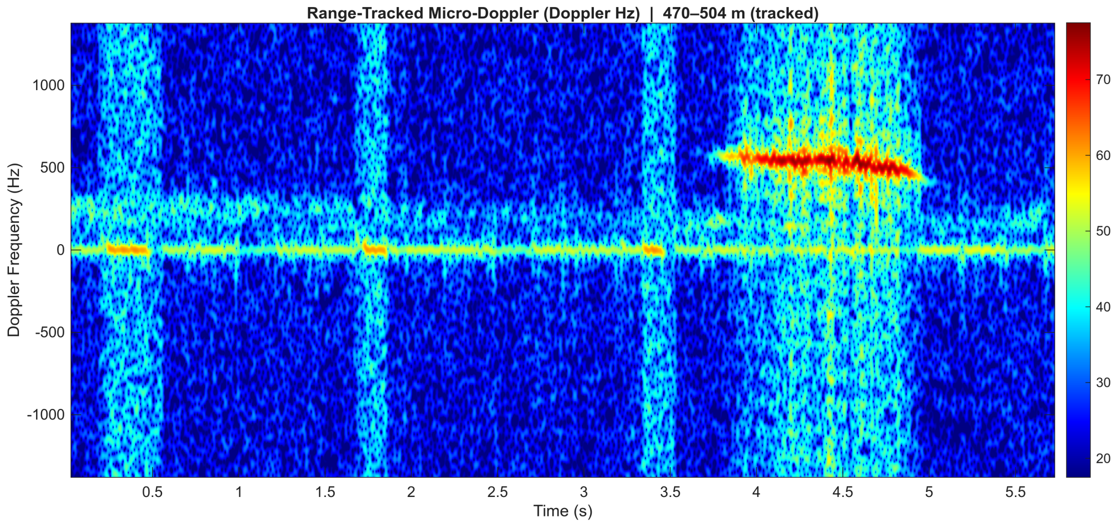


Figure 5.6: Representative drone result from the Galveston collections. The single panel shows a range-tracked Doppler-frequency spectrogram for a small rotary-wing UAS observed along the seawall at low elevation over the 470 m to 504 m tracked gate.

5.5.1 Interpretation of Representative Time–Frequency Features

For the narrowband Doppler interpretation used here, radial velocity and Doppler frequency are related by

$$f_d = \frac{2v_r}{\lambda}, \qquad v_r = \frac{\lambda f_d}{2}, \tag{5.28}$$

which, for $\lambda = 0.0319\,\mathrm{m}$, gives approximately $0.0160\,\mathrm{m/s/Hz}$. The slow-time sampling rate limits the principal Doppler interval to $\pm\mathrm{PRF}/2 = \pm 1378.9\,\mathrm{Hz}$, or approximately $\pm 21.96\,\mathrm{m/s}$.

Range tracking and Doppler concentration. The tracked products contain narrow, persistent Doppler ridges together with weaker time-varying components. A narrow ridge is consistent with a dominant rigid-body or slowly varying scattering component. It is also qualitatively consistent with reduced pulse-to-pulse phase variation after COR, although ridge width alone is not a complete coherence metric. The tracked products provide within-scene examples of how the selected range gate determines the time–frequency representation; they do not constitute an aggregate tracking-accuracy measurement.

Bird-scene excursions. The representative bird products contain time-localized Doppler excursions around the persistent ridge. Such excursions are compatible with articulated wing motion reported in the micro-Doppler literature [4], [21]. The present figures support a scene-level phenomenological interpretation only: target species, wing-tip speed, and flap–glide state are not independently measured in these collections, so those quantities are not inferred as validated labels from the spectrogram alone.

Slow ridge displacement. Smooth displacement of the dominant ridge over several seconds is consistent with a change in bulk radial velocity or target–radar geometry. It should be separated conceptually from the faster modulation attributed to component motion. Without an independently measured trajectory, the specific maneuver that produced the ridge displacement cannot be determined from the spectrogram.

Drone-scene structure. In the representative seawall drone scene, energy is concentrated closer to the persistent ridge and includes faster modulation compatible with continuously rotating components. This is qualitatively different from the time-localized excursions in the selected bird scenes, but one drone scene and two bird scenes do not establish a repeatable class-separation statistic. The classifier in Chapter 6 is therefore not claimed to have learned these Furuno-specific features.

Synchrosqueezed view and Doppler folding. The Fourier synchrosqueezed transform (FSST) reassigns energy in the time–frequency plane but does not remove slow-time aliasing. Velocity must still be interpreted modulo $\lambda\,\mathrm{PRF}/2 \approx 43.9\,\mathrm{m/s}$. For example, a component at $+7\,\mathrm{m/s}$ and its wrapped representation near $-36.9\,\mathrm{m/s}$ correspond to the same sampled Doppler class. FSST ridge concentration should therefore not be interpreted as ambiguity resolution.

Broadband and edge-of-band features. Vertical features spanning much of the Doppler interval and narrow bands near the slow-time Nyquist limits are treated as possible transients, clutter contamination, leakage, or processing artifacts rather than assigned to target micro-motion without additional evidence. Quantitative descriptors should therefore be computed within a documented Doppler interval and accompanied by sensitivity analysis of the selected range gate and time–frequency parameters.

5.6 Discussion

Several points follow from the current implementation. First, the maintained baseline includes conditional residual IF-offset correction, alignment diagnostics, and range tracking before the micro-Doppler products are formed. Second, the extension is a set of mathematically motivated estimators, not a guarantee of improvement. The assumptions behind Doppler filtering, singular-value thresholding, and inter-CPI accumulation are only approximate for finite, clutter-contaminated leakage blocks. Their effects must therefore be reported for each archived comparison. Third, the workflow retains interactive range-window selection and scene-specific parameter choices, so the figures are representative case studies rather than samples from an automated aggregate evaluation.

The baseline and extension consequently support different levels of inference. The baseline documents the processing used for the primary scene products. The extension tests whether clutter-aware reference estimation or a longer phase-corrected dwell changes a particular result. A batch protocol with fixed selection rules and quantitative uncertainty analysis is required before population-level claims can be made.

5.7 Summary

The Furuno COR contribution is a two-stage framework. The baseline stage applies adaptive IF correction, leakage-referenced pulse-phase rotation, POC alignment diagnostics with pass-through timing, matched filtering, and tracked micro-Doppler analysis. The extension estimates a clutter-aware leakage reference and inter-CPI phase corrections for longer-dwell comparisons. These algorithms produce the representative Galveston products reported in this chapter. The subsequent classifier study is performed separately on DIAT-μSAT.

Chapter 6

Lightweight Hybrid Vision Transformer for Aerial Target Classification

The measured-data and simulation chapters motivate time–frequency structures relevant to aerial-target classification, but neither supplies training data for the classifier evaluated here. Chapter 5 presents COR-processed Furuno products, and Chapter 3 examines how modeled component motion and viewing geometry affect simulated signatures. This chapter presents a separate benchmark study using a **FAN-integrated Lightweight Hybrid Vision Transformer** (FAN-LH-ViT). The architecture combines multiscale convolutional feature extraction, squeeze-and-excitation channel recalibration, transformer-style self-attention, and Fourier Analysis Network (FAN) layers [13], [14].

The architecture is motivated by periodic and quasi-periodic modulation produced by wings, rotors, and propellers [4]–[6]. Convolution supplies a local receptive-field bias, whereas self-attention permits direct interaction between nonadjacent feature tokens. FAN sine–cosine projections are inserted before self-attention to supply periodic basis components to the token representation. An ablation isolating their contribution is not reported.

The experiments reported in this chapter use the public DIAT-μSAT dataset introduced by Kumawat et al. [9]. This dataset contains labeled small-aerial-target radar recordings for six classes. In the present implementation, each MATLAB sample is loaded from its baseband signal and sampling-rate fields, converted into a two-sided micro-Doppler spectrogram, normalized to a fixed 256×256 single-channel tensor, and used to train the proposed model with categorical cross-entropy, Adam optimization, and early stopping. In addition to classification performance, the implementation records the train/validation split, training history, retained checkpoint, confusion matrix, and model parameter count.

This chapter formalizes the classification problem, presents the implemented architecture, explains the role of FAN-integrated attention in modeling periodic radar signatures, and situates the proposed network within the broader context of micro-Doppler-based aerial target recognition.

6.1 Introduction

For small airborne objects, bulk radial velocity and coarse radar cross section may not provide sufficient class separation [1]–[3]. Micro-Doppler signatures provide complementary measurements of rotating, flapping, or articulated components [4], [5]. Prior studies have therefore examined bird–drone discrimination using Doppler and micro-Doppler features [6], [28], [33], [34].

The Furuno X-band marine radar used in the measured chapters is a magnetron system and does not provide an inherent pulse-to-pulse phase reference [35]–[37]. Chapter 5 uses sampled transmit leakage, adaptive IF tracking, and pulse-wise phase rotation to form representative range–Doppler and micro-Doppler products [7]. The classifier experiment reported here is evaluated on the public DIAT-μSAT dataset and is not a classification test on those Furuno products.

In the time–frequency domain, aerial-target classification can be formulated as supervised classification of structured arrays. Deep learning methods, including convolutional neural networks and generative models, have been applied to radar micro-Doppler classification and denoising [11], [12], [29], [30]. Micro-Doppler spectrograms are not generic natural images: the vertical axis represents Doppler frequency and the horizontal axis represents slow time. Spatial shifts and resizing along these axes alter signal-domain coordinates [5], [38]. Furthermore, important class-defining patterns may span long temporal spans. For example, periodic rotor signatures from drones may recur across multiple regions of the spectrogram. At the same time, bird wingbeats often generate repeated arcs and alternating energy concentrations over time [6], [21], [39], [40]. A successful classifier must therefore capture both local time–frequency texture and long-range temporal relationships.

The simulations in Chapter 3 illustrate how viewing geometry and multi-target superposition can change time–frequency structure. These effects motivate attention to nonlocal spectrogram relationships, but the DIAT-μSAT experiment does not isolate aspect angle or flock formation as controlled factors. Generalization across those variables is therefore a future validation question rather than a demonstrated property of the architecture.

Classical convolutional neural networks are naturally suited to local feature extraction, but their receptive fields remain localized unless considerable depth is used. Vision Transformers overcome this limitation through self-attention, enabling direct interactions between distant image regions. Nevertheless, standard transformer architectures are computationally expensive and often require substantial training data. To address these competing considerations, this chapter adopts a *hybrid* approach: a lightweight pyramid of residual squeeze-and-excitation blocks first extracts compact radar features, followed by transformer-inspired attention blocks that model global dependencies [13].

The principal modification beyond a standard lightweight hybrid Vision Transformer is the insertion of a **Fourier Analysis Network (FAN) layer** [14] into the token-processing pathway of each Radar-ViT block. This insertion is motivated by micro-Doppler periodicity: the signatures are shaped by periodic or quasi-periodic target motions; therefore, explicitly encoding tokens via sine and cosine feature projections before attention yields a periodicity-aligned latent representation. The resulting model combines depthwise-separable convolution, self-attention, and FAN tokenization.

6.2 Problem Formulation

Let $x[n]$ denote the discrete-time radar signal loaded from a DIAT-μSAT sample after mean removal and, when required, resampling to the target sampling rate used by the implementation. The time–frequency structure of the signal is represented via the short-time Fourier transform (STFT),

$$X(m,k) = \sum_{n=-\infty}^{\infty} x[n]\, w[n-m] e^{-j2\pi kn/N}, \tag{6.1}$$

where $w[\cdot]$ is the Hann analysis window, m indexes slow-time frames, k is the Doppler-frequency bin index, and N is the FFT size [2], [16]. In the reported implementation, $N = 512$, the analysis window length is 256 samples, and adjacent windows overlap by 240 samples.

The spectrogram magnitude in logarithmic scale is defined as

$$S_{\mathrm{dB}}(m,k) = 20 \log_{10}\Big(|X(m,k)| + \epsilon\Big), \tag{6.2}$$

where ϵ is a small positive constant used to avoid singularity at zero magnitude. In the implemented code, $\epsilon = 10^{-12}$. The spectrogram is shifted so that its peak is at 0 dB, clipped to a 60 dB dynamic range, normalized to $[0, 1]$, and FFT-shifted so that the zero-Doppler component lies at the center of the frequency axis.

After interpolation to the fixed image grid, each sample is represented by a fixed-size single-channel tensor

$$\mathbf{X} \in \mathbb{R}^{1\times H\times W}, \qquad H = 256,\ W = 256. \tag{6.3}$$

The classification model learns a mapping

$$f_\theta : \mathbb{R}^{1\times 256\times 256} \to \mathbb{R}^C, \tag{6.4}$$

where $C = 6$ is the number of classes and θ denotes the trainable parameters.

The output logits are converted to class-probability estimates through the softmax function:

$$p(y = c \mid \mathbf{X}) = \frac{\exp(z_c)}{\sum_{j=1}^{C} \exp(z_j)}, \qquad c = 1, \ldots, C, \tag{6.5}$$

where $\mathbf{z} = f_\theta(\mathbf{X})$.

6.3 Data Preparation and Spectrogram Generation

6.3.1 Dataset Provenance and Processing Flow

The learning workflow begins with the public DIAT-μSAT dataset [9], which provides labeled samples for six aerial-target categories. In contrast to the Furuno chapters, the starting point here is not a measured Furuno intermediate-frequency recording. The implementation loads each benchmark MATLAB file, validates the required sampling-rate and signal fields, and computes the micro-Doppler image used for classification. Figure 6.1 summarizes the resulting preprocessing and training flow used in this chapter.

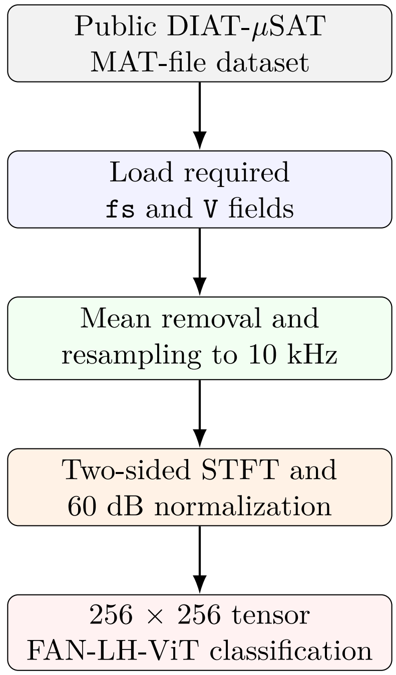


Figure 6.1: Benchmark-data processing flow used in this chapter. The classifier operates on normalized micro-Doppler images generated from DIAT-μSAT MATLAB samples rather than on Furuno IF recordings.

6.3.2 Signal Preprocessing

The public benchmark used in this chapter is processed from MATLAB files containing a sampling rate and signal vector. The implementation requires the fields `fs` and `V`. Each signal is converted to a one-dimensional floating-point sequence, mean-centerd, and resampled to 10 kHz when the original sampling rate is higher than that target rate. Specifically, the complex-valued signal vector `V` is resampled in time; the spectrogram and the final image are not resampled at this stage. A two-sided STFT is then computed using a Hann window with $M = 256$, overlap 240, and $N_{\mathrm{FFT}} = 512$. The resulting log-magnitude spectrogram is peak-normalized, clipped to a 60 dB dynamic range, mapped to $[0, 1]$, FFT-shifted along the frequency axis, and interpolated to 256×256 pixels. This processing produces the tensor $\mathbf{X} \in \mathbb{R}^{1\times256\times256}$ used by the network.

6.3.3 Class Structure

The classification problem follows the six folder labels provided by the DIAT-μSAT benchmark. Table 6.1 reports the stratified train/validation split used by the run reported in this chapter. The labels are retained as benchmark labels; no additional class semantics are inferred from local example-image filenames.

Table 6.1: DIAT-μSAT class labels and stratified split counts used for the reported FAN-LH-ViT run.

Dataset label	Training samples	Validation samples
`3_blades_Short`	560	240
`3_blades_long`	559	240
`Birdminihelicopter`	571	244
`RC-Plane`	560	240
`bird`	560	240
`drone`	584	251
Total	**3,394**	**1,455**

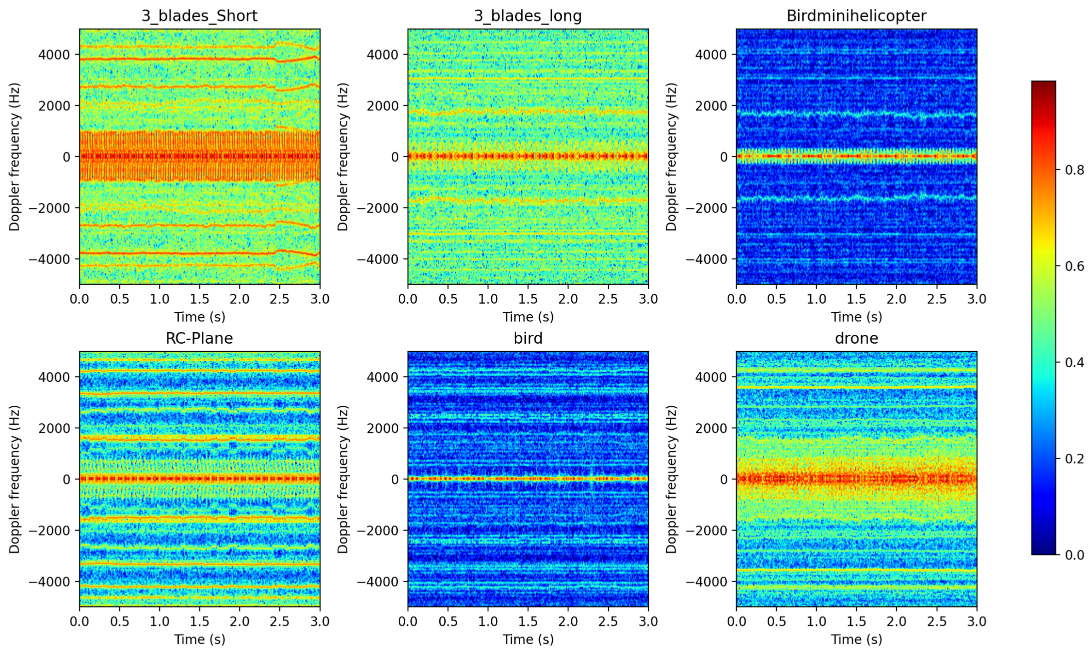


Figure 6.2: Representative normalized micro-Doppler spectrograms generated by the corrected DIAT-μSAT preprocessing code for the six benchmark labels.

The class structure contains both coarse categories and fine-grained benchmark labels. Distinguishing birds from rotorcraft-like signatures and separating the two three-blade labels requires sensitivity to periodic modulation structure, which motivates the use of FAN-based token encoding. Because these categories come from a published benchmark, they should be interpreted as the taxonomy of the DIAT-μSAT dataset rather than as a restatement of the Galveston scene folders discussed elsewhere in the thesis.

6.4 FAN-Integrated LH-ViT Architecture

The implemented network is a FAN-integrated lightweight hybrid Vision Transformer based on the LH-ViT(4-2-1) macro-architecture [13]. FAN layers are inserted before self-attention in the two Radar-ViT blocks. The remaining components are a four-stage residual squeeze-and-excitation (RES-SE) feature pyramid, one auxiliary RES-SE block, global average pooling, dropout, and a linear classification head. This hybrid design is intended to exploit the strengths of both convolution and attention:

1) Convolution extracts local and multiscale time–frequency features.

2) Self-attention relates nonadjacent feature tokens.

3) FAN layers introduce sine–cosine projections before attention.

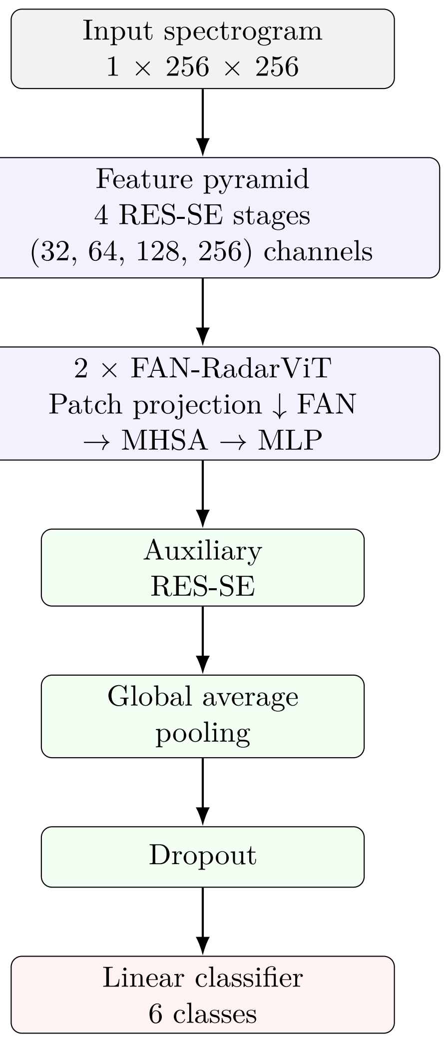


Figure 6.3: Architecture of the FAN-LH-ViT classifier.

Table 6.2: Architecture specification of the proposed FAN-LH-ViT model.

Stage	Module	Output channels	Function
Input	Normalized spectrogram	1	Single-channel STFT magnitude map
Pyramid 1	2 × RES-SE blocks	32	Local feature extraction and downsampling
Pyramid 2	2 × RES-SE blocks	64	Mid-level feature extraction
Pyramid 3	2 × RES-SE blocks	128	Hierarchical abstraction
Pyramid 4	2 × RES-SE blocks	256	Deep compact feature representation
Attention 1	FAN-RadarViT block	256	FAN tokenization, multi-head self-attention (MHSA), and MLP
Attention 2	FAN-RadarViT block	256	Global context refinement
Refinement	Auxiliary RES-SE block	256	Post-attention refinement
Head	Global average pooling, dropout, linear layer	6	Classification logits

Table 6.3: Training and model hyperparameters used for the reported FAN-LH-ViT run.

Parameter	**Value**
Input size	$1 \times 256 \times 256$
Number of classes	6
Dataset size	4,849 samples
Split policy	Stratified 70/30 hold-out, random seed 32
Base channels	32
Number of pyramid levels	4
Number of FAN-RadarViT blocks	2
Patch size	2
Number of attention heads	4
MLP expansion ratio	2
FAN periodic ratio p_{ratio}	0.30
Dropout probability	0.1
Batch size	16
Learning rate	10^{-4}
Optimizer	Adam
Weight decay	10^{-4}
Maximum epochs	35
Validation split	0.3
Early stopping patience	10
Early stopping delta	0.01

6.4.1 Squeeze-and-Excitation Block

The squeeze-and-excitation (SE) block adaptively rescales feature channels. For an input tensor

$$\mathbf{X} \in \mathbb{R}^{B\times C\times H\times W}, \tag{6.6}$$

global average pooling first computes a channel descriptor:

$$s_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} X_{c,i,j}. \tag{6.7}$$

This descriptor is passed through a bottleneck multilayer perceptron:

$$\mathbf{a} = \sigma\Big(\mathbf{W}_2\,\delta(\mathbf{W}_1\mathbf{s})\Big), \tag{6.8}$$

where $\delta(u) = \max(0, u)$ is the ReLU activation and $\sigma(u) = 1/(1 + e^{-u})$ is the sigmoid function [41]. The matrix $\mathbf{W}_1$ is the learned channel-reduction projection in the bottleneck, and $\mathbf{W}_2$ is the learned channel-expansion projection that returns the descriptor to C channel weights. The output is then

$$\tilde{X}_{c,i,j} = a_c X_{c,i,j}. \tag{6.9}$$

This mechanism provides input-dependent channel reweighting. Its use in the present model is inherited from the lightweight hybrid architecture [13]; the contribution of the squeeze-and-excitation blocks is not isolated experimentally here.

6.4.2 Depthwise Separable Convolution

Depthwise separable convolution reduces computational complexity by decomposing standard convolution into depthwise and pointwise operations. If the input contains C channels, X_c denotes the cth input feature map, $K_c^{(d)}$ is its depthwise spatial kernel, $*$ denotes spatial convolution, and Y_c is the resulting depthwise feature map:

$$Y_c = X_c * K_c^{(d)}. \tag{6.10}$$

A pointwise 1×1 convolution then mixes these channels. Here, $W_{m,c}^{(p)}$ is the learned pointwise weight from input channel c to output channel m, and Z_m is the mth output feature map:

$$Z_m = \sum_{c=1}^{C} W_{m,c}^{(p)} Y_c. \tag{6.11}$$

This factorization substantially reduces parameter count and floating-point operations relative to full convolution, which is desirable in lightweight radar classification networks [13].

6.4.3 RES-SE Block

Each RES-SE block contains a 1×1 projection convolution, batch normalization, ReLU activation, a depthwise separable convolution, an SE block, and a residual shortcut. These are the projection, normalization, activation, depthwise-separable, channel-recalibration, and residual components of the lightweight hybrid design described in the cited architecture and component references [13], [41], [42].

If $\mathcal{F}(\mathbf{X})$ denotes the nonlinear branch and $\mathcal{S}(\mathbf{X})$ the shortcut branch, the output is

$$\mathbf{Y} = \mathrm{ReLU}\Big(\mathcal{F}(\mathbf{X}) + \mathcal{S}(\mathbf{X})\Big). \tag{6.12}$$

Residual learning stabilizes the optimization process and enables deeper hierarchical extraction of radar features.

6.4.4 Feature Pyramid

The feature pyramid contains four stages. Starting from a single input channel, the number of output channels doubles across stages:

$$1 \rightarrow 32 \rightarrow 64 \rightarrow 128 \rightarrow 256. \tag{6.13}$$

Consequently, the spatial dimensions are gradually reduced while the semantic richness of the feature maps increases. For an input of size 256×256, the effective spatial progression is approximately

$$256 \times 256 \rightarrow 128 \times 128 \rightarrow 64 \times 64 \rightarrow 32 \times 32 \rightarrow 16 \times 16. \tag{6.14}$$

The pyramid provides feature maps at successively coarser spatial scales before attention is applied. The channel progression and stage roles are summarized in Table 6.2.

6.4.5 Fourier Analysis Network Layer

The principal architectural modification is the FAN layer inserted inside each Radar-ViT block. It maps an input vector $\mathbf{x} \in \mathbb{R}^d$ to a representation containing periodic and nonperiodic branches [14]:

$$\phi(\mathbf{x}) = \Big[\cos(\mathbf{W}_p\mathbf{x} + \mathbf{b}_p) \parallel \sin(\mathbf{W}_p\mathbf{x} + \mathbf{b}_p) \parallel g(\mathbf{W}_g\mathbf{x} + \mathbf{b}_g)\Big], \tag{6.15}$$

where:

- $\mathbf{W}_p$ is the learned projection from the d-dimensional token into the periodic subspace; its rows determine the learned harmonic directions,
- $\cos(\cdot)$ and $\sin(\cdot)$ provide explicit harmonic basis functions,
- $\mathbf{W}_g$ is the learned projection into the complementary nonperiodic feature branch,
- $g(\cdot)$ is the Gaussian Error Linear Unit (GELU) activation.

The periodic branch dimension is controlled by $p_{\text{ratio}} = 0.30$, so the dimensions of $\mathbf{W}_p$ and $\mathbf{W}_g$ are selected so that the concatenated output has the token embedding dimension C at that stage. The biases $\mathbf{b}_p$ and $\mathbf{b}_g$ are learned offsets. GELU is defined by $g(u) = u\Phi(u)$, where Φ is the standard normal cumulative distribution function, and is commonly approximated by $g(u) \approx \frac{1}{2}u[1+\tanh(\sqrt{2/\pi}(u+0.044715u^3))]$ [14]. The sine and cosine branches provide bounded periodic features, while the GELU branch supplies a smooth nonlinear complement.

Periodic component motion is a source of micro-Doppler modulation [4], [5]. Rotor blades, wingbeats, and other oscillatory mechanisms can generate harmonic or quasi-harmonic time–frequency structure [6], [21], [43]. The sine and cosine branches therefore provide a radar-motivated architectural bias; the reported hold-out experiment does not prove that this branch causes the observed accuracy.

6.4.6 FAN-RadarViT Block

The FAN-RadarViT block modifies a standard lightweight Radar-ViT block by inserting the FAN layer before the self-attention layer.

Let the incoming feature map be

$$\mathbf{X} \in \mathbb{R}^{B\times C\times H\times W}. \tag{6.16}$$

Patch projection is first performed using a convolution with kernel size and stride equal to the patch size $P = 2$:

$$\mathbf{X}_{\text{proj}} = \text{Conv2D}_{P,P}(\mathbf{X}). \tag{6.17}$$

The projected feature map is flattened into a token sequence

$$\mathbf{T} \in \mathbb{R}^{B\times N\times C}, \qquad N = \frac{H}{P}\frac{W}{P}. \tag{6.18}$$

The FAN layer then processes each token. If $\mathbf{t}_{b,n} \in \mathbb{R}^C$ denotes the nth token of batch item b, then

$$\tilde{\mathbf{t}}_{b,n} = \phi(\mathbf{t}_{b,n}), \tag{6.19}$$

where $\phi(\cdot)$ is defined in (6.15). These FAN-transformed tokens are then normalized:

$$\mathbf{T}_1 = \mathrm{LN}(\tilde{\mathbf{T}}), \tag{6.20}$$

and passed through multi-head self-attention. For attention head i, the normalized token matrix is projected into queries, keys, and values:

$$\mathbf{Q}_i = \mathbf{T}_1\mathbf{W}_i^Q, \qquad \mathbf{K}_i = \mathbf{T}_1\mathbf{W}_i^K, \qquad \mathbf{V}_i = \mathbf{T}_1\mathbf{W}_i^V,$$

where $\mathbf{W}_i^Q$, $\mathbf{W}_i^K$, and $\mathbf{W}_i^V$ are learned projection matrices. The query vectors represent the tokens seeking information, the key vectors describe the tokens being compared, and the value vectors contain the feature content that is mixed. If there are N tokens, then $\mathbf{Q}_i$ and $\mathbf{K}_i$ have N rows, and the matrix $\mathbf{Q}_i\mathbf{K}_i^\top$ contains the pairwise token similarities.

$$\mathrm{MHSA}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathrm{Concat}(\mathrm{head}_1, \ldots, \mathrm{head}_h)\mathbf{W}_O, \tag{6.21}$$

with

$$\mathrm{head}_i = \mathrm{softmax}\left(\frac{\mathbf{Q}_i\mathbf{K}_i^\top}{\sqrt{d_k}}\right)\mathbf{V}_i. \tag{6.22}$$

The softmax converts each row of the scaled similarity matrix into nonnegative attention weights that sum to one. Thus, each output token is a weighted combination of the value vectors, with larger weights assigned to more similar query–key pairs [44].

A residual attention update is applied:

$$\mathbf{T}_2 = \tilde{\mathbf{T}} + \text{Dropout}\Big(\text{MHSA}(\mathbf{Q}, \mathbf{K}, \mathbf{V})\Big), \tag{6.23}$$

followed by a residual MLP update:

$$\mathbf{T}_3 = \mathbf{T}_2 + \text{Dropout}\Big(\text{MLP}(\text{LN}(\mathbf{T}_2))\Big). \tag{6.24}$$

Finally, the tokens are folded back to a spatial feature map and bilinearly upsampled:

$$\mathbf{Y} = \text{Interp}\Big(\text{Fold}(\mathbf{T}_3)\Big). \tag{6.25}$$

This block combines patch-based contextual interaction, explicit harmonic token encoding, and lightweight computation. The output token matrix $\mathbf{T}_3$ is folded back into the spatial arrangement of the projected feature map, and $\text{Interp}(\cdot)$ restores the spatial resolution used by the following convolutional stage.

6.4.7 Classifier Head

After two FAN-RadarViT blocks, an auxiliary RES-SE block refines the deep feature tensor. Global average pooling then collapses the spatial dimensions:

$$\mathbf{z}_{\text{gap}} = \text{GAP}(\mathbf{Y}) \in \mathbb{R}^{C_f}, \tag{6.26}$$

where C_f is the final channel count and $\text{GAP}(\cdot)$ averages each channel over its spatial positions. Dropout randomly masks a subset of these features during training to reduce co-adaptation. The retained feature vector is then passed to the final linear classifier:

$$\mathbf{z} = \mathbf{W}\ \mathrm{Dropout}(\mathbf{z}_{\mathrm{gap}}) + \mathbf{b}. \tag{6.27}$$

where $\mathbf{W}$ and $\mathbf{b}$ are the learned classifier weights and bias. The resulting logits $\mathbf{z} \in \mathbb{R}^6$ are converted to six class-probability estimates by the softmax function.

6.5 Training Methodology

6.5.1 Dataset Split

The experimental configuration uses a stratified hold-out split of the DIAT-μSAT benchmark with a validation fraction of 0.3, so that approximately 70% of samples are used for training and 30% for hold-out validation. The split is generated with random seed 32 and stratification over the six class labels. The resulting split contains 3,394 training samples and 1,455 validation samples, with the class counts reported in Table 6.1.

This split is adequate for a proof-of-concept model-development experiment on the public benchmark, but it is not equivalent to validation on independent Furuno field collections. The results reported in this chapter should therefore be interpreted as hold-out performance on the published benchmark rather than as a definitive estimate of generalization across independent field collections. Source-record or acquisition-group disjointness has not been verified beyond the published benchmark labels.

6.5.2 Optimization

The main optimization settings are a scheduled maximum of 35 epochs, batch size 16, learning rate 10^{-4}, Adam optimization, and weight decay 10^{-4}. The objective function is the multi-class cross-entropy loss:

$$\mathcal{L}_{\mathrm{CE}} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{\exp\left(z_{y^{(i)}}^{(i)}\right)}{\sum_{c=1}^{C}\exp\left(z_c^{(i)}\right)}. \tag{6.28}$$

6.5.3 Early Stopping

To mitigate overfitting, early stopping is used with a patience of 10 epochs and a minimum improvement threshold of $\delta = 0.01$. Let the validation loss at epoch e be $\mathcal{L}_{\mathrm{val}}^{(e)}$. Training is terminated if the validation loss fails to improve by at least δ over 10 successive epochs. The best model state under this delta-qualified rule is retained. In the reported run, training stops at epoch 30 and reloads the retained epoch-20 model state before final reporting.

6.5.4 Evaluation Metrics

At each epoch, the implementation computes the training loss, validation loss, accuracy, weighted precision, weighted recall, weighted F1-score, and a full per-class classification report. These metrics provide a more complete assessment than accuracy alone, especially when class frequencies are not perfectly uniform.

6.6 Experimental Results

The performance of the FAN-LH-ViT model is summarized in Table 6.4. The reported run is the corrected FAN-LH-ViT training run archived as `training_log_20260430_141118.txt`, `history_20260430_141118.json`, and `classification_-`

`report_20260430_141118.txt`. Under the configured early-stopping rule, the retained checkpoint is selected at epoch 20, with validation loss 0.0457, accuracy 98.63%, and weighted precision, recall, and F1-score all equal to 0.9863. The held-out validation subset contains 1,455 samples distributed across the six stratified DIAT-μSAT labels.

The epoch-30 logged validation loss is lower in raw value (0.0403), but the early-stopping object retains model states only when the validation-loss improvement exceeds $\delta = 0.01$. The final classification report therefore evaluates the retained epoch-20 model state. These results should be interpreted as benchmark evidence for the proposed classifier rather than as direct validation on the measured Furuno or Galveston collections.

Table 6.4: Classification performance of the FAN-LH-ViT model at the retained best checkpoint under the early-stopping rule (epoch 20, $p_{\text{ratio}} = 0.3$).

Class	**Precision**	**Recall**	**F1-score**	**Support**
`3_blades_Short`	0.9915	0.9750	0.9832	240
`3_blades_long`	0.9792	0.9792	0.9792	240
`Birdminihelicopter`	0.9718	0.9877	0.9797	244
`RC-Plane`	0.9794	0.9917	0.9855	240
`bird`	0.9958	0.9875	0.9916	240
`drone`	1.0000	0.9960	0.9980	251
Accuracy		0.9863		1,455
Macro average	0.9863	0.9862	0.9862	1,455
Weighted average	0.9863	0.9863	0.9863	1,455

Figure 6.4: Training and validation loss curves for the corrected FAN-LH-ViT run. The final report reloads the retained epoch-20 checkpoint selected by the early-stopping rule.

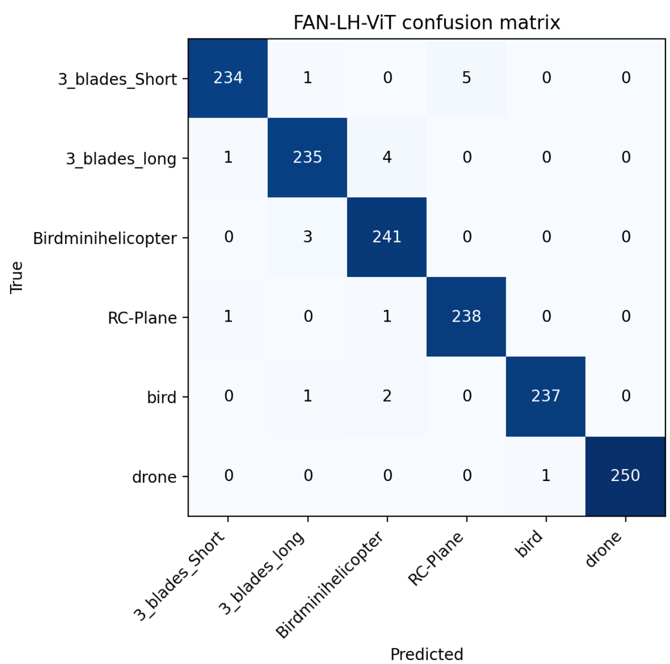


Figure 6.5: Confusion matrix for the retained FAN-LH-ViT checkpoint on the 1,455-sample stratified validation split. Rows correspond to true labels and columns correspond to predicted labels.

The closest published point of reference is the work of Kumawat et al. [9], who introduced the public DIAT-μSAT dataset for small-aerial-target micro-Doppler classification. The present run uses the same public dataset family but a separately implemented STFT pipeline, stratified split, and FAN-LH-ViT model. A controlled accuracy comparison against the original CNN/VGG experiments is not claimed here because the original preprocessing and split policy were not rerun in this thesis implementation.

For a structural breakdown, Table 6.5 quantifies the parameter-count gap between grayscale-adapted VGG models and the proposed model. Simonyan and Zisserman introduced VGG as a sequence of 3×3 convolutional blocks terminated by three fully connected layers with 4,096 channels each [45]. In this thesis comparison, the first VGG convolution is adapted to a single-channel input to match radar spectrograms. The resulting VGG16 and VGG19 parameter counts are approximately 58× and 60× larger than the FAN-LH-ViT count computed from the current PyTorch model definition.

Table 6.5: Parameter-count comparison between grayscale-adapted VGG models [45] and the proposed FAN-LH-ViT model. Parameter memory assumes 32-bit floating-point storage and does not include optimizer states or activation memory.

Model	**Trainable params.**	**Param. memory (MB)**	**Relative params.**
VGG16, grayscale input [45]	138 356 392	527.79	58.2×
VGG19, grayscale input [45]	143 666 088	548.04	60.5×
FAN-LH-ViT (proposed)	2 375 918	9.06	1.0×

Three structural differences motivate the FAN-LH-ViT design despite its smaller parameter count. These are design motivations rather than ablation-proven causal claims.

First, VGG employs a stack of dense 3×3 convolutions and does not include squeeze-and-excitation channel reweighting. FAN-LH-ViT includes input-dependent inter-channel weights. This is a structural distinction, not evidence that a particular channel corresponds to rotor Doppler or clutter.

Second, VGG processes inputs through a strictly local convolutional hierarchy, so long-range temporal relationships in a time–frequency spectrogram—such as the recurrence period of a drone's rotor modulation—are represented only after several convolution and pooling stages. FAN-LH-ViT incorporates transformer self-attention after convolutional feature compression, allowing non-adjacent spectrogram regions to interact within the token sequence [13].

Third, VGG contains no explicit inductive bias for oscillatory structure. The FAN layers replace generic linear token projections with sine–cosine feature mappings, encoding each spectrogram patch in a Fourier-compatible representation before self-attention measures inter-token similarity. This architectural choice is motivated by the periodic structure of radar micro-Doppler, where class-discriminating patterns arise from rotating or flapping motions [4], [14]. Training FAN-LH-ViT from scratch on the radar-domain tensors also removes dependence on ImageNet initialization, although no scratch-versus-transfer ablation is reported here.

6.7 Computational Considerations

A practical motivation for the proposed architecture is parameter efficiency. The current FAN-LH-ViT implementation has 2 375 918 trainable parameters for six-class classification on $1 \times 256 \times 256$ tensors. This is much smaller than grayscale-adapted VGG16 and VGG19 parameter counts, as summarized in Table 6.5.

This is important because transformer-inspired radar models are often criticized for excessive computational overhead. The use of depthwise separable convolutions, pyramidal downsampling, only two attention blocks, and a compact classifier head keeps the network small enough to motivate future embedded evaluation [13], [46].

The parameter comparison against VGG16 and VGG19 adapted for grayscale input highlights the trade-off between model capacity and computational burden. Forward/backward activation memory is not reported as a thesis result because it depends on batch size, input resolution, device, and summary-tool implementation. The bounded claim supported here is that the proposed architecture attains high hold-out accuracy on DIAT-μSAT with far fewer trainable parameters than VGG-style baselines.

6.8 Scope and Threats to Validity

The classifier results in this chapter should be interpreted within the scope of the public benchmark used here. The strongest measured-data evidentiary chain in the thesis runs from Galveston IF collections through coherent-on-receive processing to the resulting micro-Doppler products. By contrast, the classification stage reported in this chapter is a benchmark study on DIAT-μSAT rather than a train-and-test experiment on the measured Furuno collections.

Two limitations follow from this. First, strong performance on DIAT-μSAT does not by itself establish generalization to the Furuno/Galveston sensing chain, and the present split has not been audited for source-record or acquisition-group disjointness. Second, comparisons with previously reported CNN or VGG-style baselines are benchmark-adjacent rather than a fully controlled rerun under identical preprocessing and split conditions. These limitations bound the result to what is supported here: a compact, periodicity-aware hybrid transformer reaches high hold-out accuracy on the public DIAT-μSAT benchmark.

6.9 Rationale for FAN Integration

The insertion of FAN layers into the Radar-ViT blocks is motivated by motion-generated periodic structure in micro-Doppler spectrograms. A standard attention block receives token embeddings formed from learned projections; no explicit sine–cosine branch is required by that formulation.

The FAN layer augments the token mapping with learned sine and cosine projections before self-attention. This supplies basis functions that can represent oscillatory variation, together with a nonperiodic nonlinear branch.

This is relevant for:

- Multi-rotor drones, which generate repeating rotor-induced micro-Doppler patterns [6], [38],
- Birds, whose wingbeat signatures are periodic but less rigidly harmonic [21], [47]. The aspect-angle and flock-diversity analyses in Chapter 3 show that bird spectrograms exhibit substantial intra-class variability driven by viewing geometry and formation effects, making periodicity-aware encoding a plausible design choice under that variability,
- Propeller-driven aircraft, whose signatures often combine steady translation with periodic propeller modulation.

The resulting FAN-RadarViT block is therefore a hybrid token-mapping and attention module motivated by micro-Doppler periodicity. The present experiment evaluates the complete model; an ablation isolating the FAN contribution remains necessary before assigning a performance gain to this component.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis examined aerial-target micro-Doppler through three separate studies: leakage-referenced processing of Furuno magnetron radar data, kinematic and scattering simulation of bird and drone signatures, and benchmark classification with a compact hybrid transformer. The evidence supports conclusions within each study, but it does not support an end-to-end Furuno classification claim.

7.1.1 Coherent-on-Receive Processing

The Furuno contribution is a layered coherent-on-receive (COR) framework for pulse-indexed analytic IF data. The maintained baseline performs adaptive IF estimation, conditional residual IF-offset correction, leakage-window selection, rank-1 SVD reference estimation, pass-through timing with phase-only-correlation diagnostics, pulse-wise leakage phase rotation, matched filtering, and range-tracked time–frequency analysis. The signal model shows that leakage rotation removes the pulse-common magnetron phase while retaining target propagation phase relative to the leakage path.

Representative Galveston bird and drone scenes contain range-localized range–

Doppler and STFT structure after processing. These results are case studies rather than an automated benchmark: range initialization and several parameters remain interactive, and aggregate phase-stability, tracking-error, and spectrogram metrics are not reported across the full collection. The clutter-aware leakage estimator and inter-CPI phase tracker are therefore treated as research extensions whose effects must be assessed scene by scene.

7.1.2 Target-Signature Simulation

The simulation study extends a two-segment wing baseline to a three-segment kinematic chain with asymmetric stroke timing, lead–lag motion, stroke-dependent orientation, body pitch and heave, and optional intermittent-flight gating. Aspect-dependent ellipsoidal scattering and Reynolds-rule agent trajectories permit controlled examination of viewing geometry and multi-bird superposition. Under the selected parameters, wing-component velocity, bulk translation, aspect, and per-bird trajectory variation produce distinct changes in the simulated Doppler ridges.

The biological scope is limited. The Herring Gull values provide a reference for selected parameters, but the simulated span and body length exceed the tabulated morphometrics, and the model is not fitted to measured three-dimensional wing trajectories or calibrated RCS data. The results consequently support kinematic interpretation and sensitivity analysis, not species-level validation or quantitative agreement with the Furuno bird scenes.

7.1.3 DIAT-μSAT Classification

The FAN-integrated Lightweight Hybrid Vision Transformer (FAN-LH-ViT) combines a residual squeeze-and-excitation convolutional pyramid, two self-attention blocks, and FAN sine–cosine token projections. On the stratified DIAT-μSAT split used in this thesis, the retained epoch-20 checkpoint attains 98.63% accuracy on 1,455 hold-out samples. The implemented model contains 2,375,918 trainable parameters, substantially fewer than the grayscale-adapted VGG parameter counts computed for the structural comparison.

This result is bounded to the reported benchmark protocol. Source-record or acquisition-group disjointness has not been independently verified, the previously published CNN/VGG pipelines were not rerun under identical preprocessing and splits, and no ablation isolates the contribution of the FAN branch. Parameter count also does not establish latency, activation memory, energy consumption, or embedded throughput. Most importantly, the classifier is not trained or evaluated on the Furuno/Galveston collections.

7.2 Future Work

1. **Automated COR evaluation:** process the archived Furuno collections with fixed target-selection and parameter rules, and report phase-error, coherence-factor, tracking-error, signal-to-clutter, and spectrogram metrics with uncertainty intervals.

2. **Furuno classification dataset:** construct source-record-disjoint training, validation, and test partitions from labeled COR-processed Furuno collections. This is required to evaluate the sensing and classification chain in a common measurement domain.

3. **Measured-model comparison:** estimate wing-beat rate, Doppler extent, ridge timing, and bulk-motion parameters from measured bird scenes and compare them with simulated distributions rather than selected example

images.

4. **Controlled architecture ablations:** compare the FAN-LH-ViT model with the same architecture without FAN projections and with convolution-only and attention-only variants under identical splits, preprocessing, and training budgets.

5. **Group-aware and external validation:** verify source-record grouping in DIAT-μSAT, repeat evaluation with group-disjoint partitions where metadata permit, and test on an independent radar dataset.

6. **Complex-valued or coherence-aware inputs:** examine whether phase-preserving representations or coherence-aware auxiliary objectives provide information beyond magnitude spectrograms [48].

7. **Implementation measurements:** report latency, peak activation memory, energy, and throughput on a specified embedded platform before making deployment claims.

The principal outcome is therefore a set of bounded contributions: leakage-referenced COR produces interpretable field-data products for representative Furuno scenes; the simulations expose how selected kinematic, scattering, and geometric variables alter modeled signatures; and FAN-LH-ViT attains high hold-out accuracy with a compact parameter count on DIAT-μSAT. Validation of the combined pipeline on measured Furuno data remains the central next step.

References

[1] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. New York, NY, USA: McGraw-Hill, 2001, ISBN: 978-0072881387.

[2] M. A. Richards, *Fundamentals of Radar Signal Processing*, 2nd ed. New York, NY, USA: McGraw–Hill Education, 2014, ISBN: 978-0071798327.

[3] S. Rahman and D. A. Robertson, "In-flight RCS measurements of drones and birds at K-band and W-band," *IET Radar, Sonar & Navigation*, vol. 12, no. 12, pp. 1493–1500, 2018. DOI: `10.1049/iet-rsn.2018.5122`.

[4] V. C. Chen, D. C. Tahmoush, and W. J. Miceli, *Radar Micro-Doppler Signatures: Processing and Applications*. The Institution of Engineering and Technology, 2014. DOI: `10.1049/PBRA034E`. eprint: `https://digital-library.theiet.org/doi/pdf/10.1049/PBRA034E`. [Online]. Available: `https://digital-library.theiet.org/doi/abs/10.1049/PBRA034E`.

[5] V. C. Chen, *The Micro-Doppler Effect in Radar*, 2nd. Artech House, 2019, ISBN: 9781608070589.

[6] S. Rahman and D. A. Robertson, "Radar micro-doppler signatures of drones and birds at K-band and W-band," *Scientific Reports*, vol. 8, p. 17 396, 2018. DOI: `10.1038/s41598-018-35880-9`.

[7] J. Horstmann, J. Bödewadt, R. Carrasco, *et al.*, "A coherent on receive X-band marine radar for ocean observations," *Sensors*, vol. 21, no. 23, p. 7828, 2021. DOI: `10.3390/s21237828`.

[8] H. Abdelbagi, M. Aljohani, A. Mrebit, *et al.*, "FPGA-based coherent doppler processor for marine radar applications," in *2015 National Aerospace and Electronics Conference (NAECON)*, IEEE, 2015, pp. 58–69. DOI: `10.1109/NAECON.2015.7443040`. [Online]. Available: `http://ieeexplore.ieee.org/document/7443040/`.

[9] H. C. Kumawat, M. Chakraborty, A. A. Bazil Raj, *et al.*, "Diat-μsat: Small aerial targets' micro-doppler signatures and their classification using cnn," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022. DOI: 10.1109/LGRS.2021.3102039.

[10] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 25–34, 1987. DOI: 10.1145/37402.37406.

[11] D. Yang, "Use of deep neural networks for classification of micro-doppler signature from radar data," Master's thesis, Chalmers University of Technology, 2018.

[12] S. Abdulatif, Q. Wei, F. Aziz, *et al.*, "Micro-doppler based human-robot classification using ensemble and deep learning approaches," in *2018 IEEE Radar Conference (RadarConf18)*, IEEE, 2018, pp. 1043–1048. DOI: 10.1109/RADAR.2018.8378705. [Online]. Available: https://ieeexplore.ieee.org/document/8378705/.

[13] S. Huan, Z. Wang, X. Wang, *et al.*, "A lightweight hybrid vision transformer network for radar-based human activity recognition," *Scientific Reports*, 2022. DOI: https://doi.org/10.1038/s41598-023-45149-5.

[14] Y. Dong, G. Li, Y. Tao, *et al.*, "FAN: Fourier analysis networks," *arXiv preprint arXiv:2410.02675*, 2024.

[15] B. R. Mahafza, *Radar Systems Analysis and Design Using MATLAB*, 4th ed. Boca Raton, FL, USA: Chapman & Hall/CRC Press, 2022, ISBN: 978-0367507930.

[16] B. R. Mahafza, *Radar Signal Analysis and Processing Using MATLAB*. Boca Raton, FL, USA: Chapman & Hall/CRC Press, 2010, ISBN: 978-1420066449.

[17] J. Metcalf, *Lecture 4: Simple Pulse and Range Resolution*, Lecture Slides, Advanced Radar Research Center, University of Oklahoma, From uploaded file: Lecture4_RangeResolution.pdf, Accessed 2024.

[18] J. Metcalf, *Lecture 5: The Doppler Shift*, Lecture Slides, Advanced Radar Research Center, University of Oklahoma, From uploaded file: Lecture5_DopplerShift.pdf, Accessed 2024.

[19] J. Metcalf, *Lecture 7: Pulsed Radar Data Structure*, Lecture Slides, Advanced Radar Research Center, University of Oklahoma, From uploaded file: Lecture7_PulsedRadarDataStructure.pdf, Accessed 2024.

[20] N. E. Huang, Z. Shen, S. R. Long, *et al.*, "The empirical mode decomposition and the hilbert spectrum for non-linear and non-stationary time series analysis," *Proceedings of the Royal Society A*, vol. 454, no. 1971, pp. 903–995, 1998. DOI: 10.1098/rspa.1998.0193.

[21] B. BRUDERER, D. PETER, A. BOLDT, *et al.*, "Wing-beat characteristics of birds recorded with tracking radar and cine camera," *Ibis*, vol. 152, no. 2, pp. 272–291, 2010. DOI: https://doi.org/10.1111/j.1474-919X.2010.01014.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1474-919X.2010.01014.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1474-919X.2010.01014.x.

[22] C. J. Pennycuick, *Modelling the Flying Bird*. Academic Press, 2008.

[23] B. BRUDERER and A. BOLDT, "Flight characteristics of birds:" *Ibis*, vol. 143, no. 2, pp. 178–204, 2001. DOI: https://doi.org/10.1111/j.1474-919X.2001.tb04475.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1474-919X.2001.tb04475.x. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1474-919X.2001.tb04475.x.

[24] B. Ramakrishnananda and K. C. Wong, "Animating bird flight using aerodynamics," *The Visual Computer*, vol. 15, no. 10, pp. 494–508, 1999. DOI: 10.1007/s003710050197.

[25] T. Weis-Fogh, "Quick estimates of flight fitness in hovering animals, including novel mechanisms for lift production," *Journal of Experimental Biology*, vol. 59, no. 1, pp. 169–230, 1973.

[26] G. Ducci, V. Colognesi, A. Vidal, *et al.*, "Stability and sensitivity analysis of bird flapping flight," *Journal of Nonlinear Science*, vol. 31, no. 2, pp. 1–30, 2021. DOI: 10.1007/s00332-020-09674-7.

[27] G. T. Ruck, D. E. Barrick, W. D. Stuart, *et al.*, *Radar Cross Section Handbook*. Plenum Press, 1970.

[28] Radio Corporation of America, "A preliminary investigation of bird classification by doppler radar," NASA, Tech. Rep. NASA CR-137457, Feb. 1973.

[29] M. Ostovan, S. Samadi, and A. Kazemi, "Generation of human micro-doppler signature based on layer-reduced deep convolutional generative adversarial network," *Computational Intelligence and Neuroscience*, vol. 2022, p. 7 365 544, 2022. DOI: `10.1155/2022/7365544`.

[30] C. Tang, W. Li, S. Vishwakarma, *et al.*, "FMNet: Latent feature-wise mapping network for cleaning up noisy micro-doppler spectrogram," *arXiv*, 2023. arXiv: `2303.00088`.

[31] Y. Ji *et al.*, "Using svd-frft filtering to suppress first-order sea clutter in hfswr," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1076–1080, May 2017. DOI: `10.1109/LGRS.2017.2697458`.

[32] M. Gavish and D. L. Donoho, "The optimal hard threshold for singular values is $4/\sqrt{3}$," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 5040–5053, 2014. DOI: `10.1109/TIT.2014.2323359`.

[33] J. Gong, J. Yan, D. Li, *et al.*, "Comparison of radar signatures based on flight morphology for large birds and small birds," *IET Radar, Sonar & Navigation*, vol. 14, no. 9, pp. 1365–1369, 2020. DOI: `10.1049/iet-rsn.2020.0064`.

[34] P. Jatau, V. Melnikov, and T.-Y. Yu, "A machine learning approach for classifying bird and insect radar echoes with S-band polarimetric weather radar," *Journal of Atmospheric and Oceanic Technology*, 2021. DOI: `10.1175/JTECH-D-20-0180.1`.

[35] FURUNO ELECTRIC CO., LTD., *Installation manual marine radar model far-1513/1523/1513-bb/1523-bb/far-1518/1528/1518-bb/1528-bb*, Pub. No. IME-36380-E, Jan. 2018.

[36] FURUNO ELECTRIC CO., LTD., *Service manual marine radar model far-1513 far-1523 far-1513-bb far-1523-bb far-1518 far-1528 far-1518-bb far-1528-bb*, Pub. No. SME-36380-C, Dec. 2019.

[37] G. Jordan, "Comparison of two major classes of coherent pulsed radar systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-11, no. 3, pp. 363–371, 1975. DOI: `10.1109/TAES.1975.308087`.

[38] K. Kim, M. Üney, and B. Mulgrew, "Estimation of drone micro-doppler signatures via track-before-detect in array radars," in *2019 International Radar Conference (RADAR)*, 2019, pp. 1–6. DOI: `10.1109/RADAR41533.2019.171375`.

[39] S. A. Gauthreaux, A.-M. Shapiro, D. Mayer, *et al.*, "Detecting bird movements with L-band avian radar and S-band dual-polarization Doppler weather radar," *Remote Sensing in Ecology and Conservation*, vol. 5, no. 3, pp. 237–246, 2019. DOI: `10.1002/rse2.101`.

[40] B. Torvik, K. E. Olsen, and H. D. Griffiths, "X-band measurements of radar signatures of large sea birds," in *International Radar Conference*, 2014.

[41] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.

[42] A. G. Howard, M. Zhu, B. Chen, *et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[43] R. Wang, C. Hu, X. Fu, *et al.*, "Micro-doppler measurement of insect wingbeat frequencies with W-band coherent radar," *Scientific Reports*, vol. 7, no. 1, p. 1396, 2017. DOI: `10.1038/s41598-017-01616-4`.

[44] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, "Attention is all you need," 2023. arXiv: `1706.03762 [cs.CL]`. [Online]. Available: `https://arxiv.org/abs/1706.03762`.

[45] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[46] J. Zabalza, C. Clemente, G. Di Caterina, *et al.*, "Robust PCA micro-doppler classification using SVM on embedded systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 2304–2310, 2014. DOI: `10.1109/TAES.2014.130082`.

[47] J. Gong, J. Yan, D. Li, *et al.*, "Using radar signatures to classify bird flight modes between flapping and gliding," *IEEE Antennas and Wireless Propagation Letters*, 2019.

[48] M. Czerkawski, C. Ilioudis, C. Clemente, *et al.*, "A novel micro-doppler coherence loss for deep learning radar applications," pp. 305–308, 2022. DOI: `10.23919/EuRAD50154.2022.9784491`.

Appendix A

Radar System and Data Acquisition Parameters

This appendix summarizes the stable radar and acquisition parameters that govern the maintained Furuno processing scripts. Earlier exploratory ASCII datasets and their scene-specific parsing tables are intentionally omitted because the current thesis narrative is grounded in the recent Galveston binary collections and the CPI-based coherent-on-receive implementation.

A.1 Furuno Radar and Acquisition Parameters

Table A.1 collects the parameters that remain fixed across the current processing workflow.

Table A.1: Stable Furuno radar and digitizer parameters used by the maintained MATLAB workflow.

Parameter	Value
Radar type	Pulsed, magnetron-based, non-coherent X-band marine radar
Carrier frequency f_c	9.41 GHz
Wavelength λ	approximately 0.0319 m
Nominal IF	approximately 60 MHz
ADC sampling interval T_s	2.5 ns
ADC sampling rate f_s	400 MHz
Short pulse used in the current scripts T_p	0.12 µs
Samples over one transmitted pulse	48
PRF	2757.8 Hz
CPI length used for COR processing	256 pulses
Sample-to-range conversion	0.375 m/sample
Nominal range resolution	approximately 18 m
Unambiguous range	approximately 54.4 km
Unambiguous velocity	approximately 21.96 m/s
Doppler-bin spacing per CPI	approximately 10.77 Hz
Velocity-bin spacing per CPI	approximately 0.172 m/s

The radar remains a magnetron system, so the above values should be understood as nominal hardware settings rather than guarantees of pulse-to-pulse coherence. In particular, the actual IF observed in the data drifts around the nominal 60 MHz center and therefore must be estimated adaptively in software.

A.2 Current Data Format and Collection Conventions

The maintained Furuno workflow now operates primarily on binary IF captures.

- **Tapping point:** The raw signal is acquired at the IF stage before envelope detection or logarithmic video processing.
- **Primary file format:** Integer binary `.bin` files. Legacy `.txt` and `.csv` readers remain in the codebase for backward compatibility, but they no

longer represent the main experimental path discussed in the thesis.

- **Filename convention:** Recent Galveston files follow the pattern `scene_name_totalSamples_samplesPerPulse.bin`. The final field is the most operationally important because it defines the pulse reshaping used by the scripts.

- **Samples per pulse:** The collections emphasized in this thesis use 8192 samples per pulse, and the COR scripts then process these pulses in blocks of 256 pulses per CPI.

- **Segmentation metadata:** Trigger location and related parsing details are handled by the MATLAB acquisition helpers and by scene-specific overrides in the Furuno scripts. These values may change with acquisition hardware or pre-processing, so they are better treated as processing metadata than as immutable thesis constants.

A.3 Representative Current Collections

Table A.2 lists representative collections used by the current archived figure set. The purpose of this table is not to claim that every file was processed identically, but to document which modern collections underpin the revised coherent-on-receive chapter.

Table A.2: Representative Galveston collections referenced by the maintained thesis figures.

Collection root	Nominal file pattern	Archived figure location
`birds_observation_pointed_north`	`*_262144_8192.bin`	`figures/galveston/...`
`birds_observation_pointed_south`	`*_262144_8192.bin`	`figures/galveston/...`

Table A.2: Representative Galveston collections referenced by the maintained thesis figures (continued).

Collection root	Nominal file pattern	Archived figure location
`seawall_drone_movement_elev0`	`*_524288_8192.bin`	`furunoReturns/figs/galveston/...`
`seawall_drone_movement_elev5_2`	`*_262144_8192.bin`	`figures/galveston/...`
`helicopter`	`*_262144_8192.bin`	`figures/galveston/...` and `furunoReturns/figs/galveston/...`

A.4 Derived Quantities Used by the Processing Chain

Several numerical quantities are derived repeatedly inside the MATLAB scripts and are listed here for convenience.

- **Pulse samples:** $N_p = \mathrm{round}(T_p/T_s) = 48$.
- **Range conversion:** $\Delta R_{\mathrm{sample}} = cT_s/2 = 0.375\,\mathrm{m}$ per sample.
- **Velocity conversion:** $f_D \mapsto v_r = \lambda f_D/2$, so $1\,\mathrm{Hz}$ corresponds to approximately $0.015\,94\,\mathrm{m/s}$.
- **CPI duration:** $T_{\mathrm{CPI}} = 256/2757.8 \approx 0.0928\,\mathrm{s}$.
- **Doppler resolution per CPI:** $\Delta f_D = \mathrm{PRF}/256 \approx 10.77\,\mathrm{Hz}$.

These are the quantities that actually anchor the current COR implementation. They govern the range axis, Doppler axis, CPI integration time, and leakage-window sizing used throughout the revised Furuno chapter.

Appendix B

Supplementary notes on Furuno Radar

The coherent-on-receive processing chain described in Chapter 5 is tightly coupled to the hardware behavior of the Furuno FAR-15x3/15x8 radar family. This appendix summarizes the system characteristics that directly motivate the implemented DSP choices and explains why the maintained CPI-based workflow takes the form it does.

B.1 Magnetron Transmitter and Non-Coherence

The Furuno FAR-15x3 and FAR-15x8 series use a magnetron as the high-power X-band source.

- **Free-running phase:** The magnetron starts each RF pulse with an effectively uncontrolled initial phase. The transmitted phase for pulse m is therefore not phase-locked to pulse $m-1$, which is the fundamental reason that conventional Doppler processing cannot be applied directly to the raw data.

- **Frequency pulling and pushing:** The service documentation specifies operation near 9.41 GHz with frequency tolerance on the order of tens of megahertz [35], [36]. In practice, the recorded IF data show both slow thermal drift and smaller pulse-to-pulse jitter, so the receiver-side DSP must estimate the effective IF adaptively rather than treat it as perfectly fixed.

- **Pulse waveform variability:** Although the transmitted pulse is nominally rectangular, the exact leakage waveform can vary slightly because of transmitter dynamics, receiver recovery, and analog filtering. This is why the matched filter is derived from the measured leakage template rather than from an idealized rectangular pulse.

B.2 Intermediate Frequency Stage and Signal Tapping

The Furuno receiver downconverts the X-band echo to a nominal IF near 60 MHz. The raw signal used in this thesis is tapped before envelope detection and before the standard marine-radar video chain. This tapping point is essential: once the signal has passed through envelope detection or logarithmic video processing, the phase information required for coherent-on-receive recovery is no longer available.

The maintained MATLAB pipeline therefore begins with adaptive IF estimation and digital downconversion. This is not merely a convenience step; it is a direct response to the fact that the nominal 60 MHz IF is only approximate for the measured data. The scripts first estimate the coarse IF from the raw spectrum, then refine it pulse by pulse, and finally apply additional gated residual CFO compensation when the measured baseband offset is large enough to warrant correction.

B.3 Triggering, Timing, and Segmentation

The external digitizer must be synchronized to the transmit trigger so that the continuous IF stream can be segmented into pulses with a consistent fast-time origin. The Furuno hardware distributes a transmit-trigger timing signal internally, and the MATLAB processing scripts rely on acquisition metadata and scene-specific parsing rules to identify the effective trigger location within each recorded pulse block.

The practical lesson is that segmentation is an acquisition-level requirement, while fine alignment must be treated as a guarded signal-processing diagnostic. Trigger synchronization brings each pulse to approximately the correct fast-time origin. The maintained pipeline retains phase-only-correlation diagnostics for residual sub-sample timing error, but the reported products use pass-through timing because enabling leakage-window shifts introduced artifacts in the current Galveston processing runs.

B.4 Transmit Leakage as the Physical Phase Reference

The coherent-on-receive method depends on the presence of a strong near-range leakage return.

- **Leakage origin:** Imperfect transmitter-receiver isolation inevitably couples a portion of the transmitted pulse into the receiver chain.
- **Fast-time location:** This leakage appears at the start of the fast-time record and is typically the strongest persistent return in the early range bins.
- **Phase relevance:** Because the leakage shares the same magnetron phase as the transmitted pulse, it is the most direct observable proxy for the otherwise unknown transmit phase. This makes it the natural reference for pulse-wise phase removal and, in the advanced script, for inter-CPI phase tracking as well.

- **Dynamic-range implications:** The leakage can be much stronger than weak far-range targets, so the data acquisition chain and downstream DSP must tolerate a very wide dynamic range. The use of SVD on the leakage window is a practical way to extract a stable reference waveform even when clutter and receiver recovery distort the raw amplitude profile.

B.5 Implications for the Maintained COR Pipeline

The hardware behavior above leads directly to the final thesis pipeline.

- The random magnetron phase requires a pulse-wise reference; the leakage provides that reference.
- IF drift and oscillator mismatch require adaptive IF tracking and gated residual CFO correction.
- Residual fast-time misregistration is monitored with phase-only-correlation diagnostics and pass-through timing in the reported runs, not by the earlier exploratory autofocus-style workflow.
- CPI-wise SVD is used because the leakage waveform is strong, low rank, and physically tied to the transmitted pulse.
- The advanced extension adds clutter-aware leakage modeling and inter-CPI phase tracking because sea clutter and long dwells can corrupt a naive single-CPI reference.

These points explain why the final implementation is organized around leakage-referenced CPI processing, matched filtering from measured templates, and range-tracked micro-Doppler extraction. The method is not an abstract algorithm layered on top of the Furuno radar; it is a software response to the exact non-idealities of that hardware platform.

Appendix C

Advanced Time-Frequency Analysis Techniques

Although the Short-Time Fourier Transform (STFT) is the time-frequency representation used throughout this thesis because of its robustness and interpretability, alternative methods are also available. Depending on the signal class, these methods can provide higher resolution or better adaptation to nonstationary behavior, often at the cost of increased computational complexity or cross-term artifacts. This appendix briefly summarizes three representative techniques: the Wigner-Ville Distribution (WVD), the Wavelet Transform (WT), and the Hilbert-Huang Transform (HHT).

C.1 Wigner-Ville Distribution (WVD)

For applications requiring higher time-frequency resolution than achievable with linear methods like STFT, bilinear transforms such as the Wigner-Ville Distribution (WVD) are employed [4], [16]. The WVD of a signal $x(t)$ is defined as the Fourier transform of its time-dependent auto-correlation function:

$$\mathrm{WVD}(t,f) = \int_{-\infty}^{\infty} x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi f\tau}d\tau \tag{C.1}$$

The WVD offers finer time-frequency resolution and energy concentration in the time-frequency plane compared to STFT, particularly for linear frequency-modulated (chirp) signals. However, a significant drawback of the WVD is its susceptibility to cross-term interference for multi-component signals. These cross-terms, which arise from the bilinear nature of the transform, appear at the mid-point in time and frequency between any two ”true” signal components. They can obscure the auto-terms and complicate the interpretation of the time-frequency representation.

To mitigate the issue of cross-terms, various smoothed or modified versions of the WVD have been developed, often belonging to Cohen's class of distributions. These include the pseudo-Wigner-Ville distribution (PWVD) and the smoothed pseudo-Wigner-Ville distribution (SPWV), which apply independent time and frequency smoothing kernels to the WVD, effectively reducing cross-term artifacts at the cost of some loss in resolution.

C.2 Wavelet Transform (WT)

The Wavelet Transform (WT) offers a multi-resolution analysis capability, providing a flexible approach to time-frequency representation that can adapt to different signal features [4], [16]. Unlike the STFT which uses a fixed-length window, the WT employs basis functions called ”wavelets,” which are scaled and translated versions of a single prototype function, the ”mother wavelet” $\psi(t)$. The Continuous Wavelet Transform (CWT) of a signal $x(t)$ is defined as:

$$\mathcal{W}_x(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a}\right) dt \tag{C.2}$$

where a is the scale parameter (inversely related to frequency) and b is the translation parameter (time shift). The scaling property is the key advantage; it allows the WT to use short windows (high frequency analysis, via small a) to achieve good time resolution for high-frequency events, and long windows (low frequency analysis, via large a) to achieve good frequency resolution for low-frequency events.

Various mother wavelets exist (e.g., Morlet, Mexican Hat, Daubechies), and the choice depends on the characteristics of the signal being analyzed. The Morlet wavelet, being a Gaussian-modulated sinusoid, is often preferred for analyzing signals with oscillatory components, such as micro-Doppler signatures, due to its good localization in both time and frequency. The squared magnitude of the CWT, $|\mathcal{W}_x(a,b)|^2$, is called a scalogram and serves as an energy density representation in the time-scale plane.

C.3 Hilbert-Huang Transform (HHT)

The Hilbert-Huang Transform (HHT) is an adaptive, data-driven time-frequency analysis technique that is well suited to nonlinear and nonstationary signals, such as those often encountered in complex micro-Doppler analysis [20]. The HHT process consists of two main parts:

1. **Empirical Mode Decomposition (EMD):** EMD adaptively decomposes a signal $x(t)$ into a finite, and often small, number of components called Intrinsic Mode Functions (IMFs), $\{c_i(t)\}$, and a residual trend $r(t)$. Each IMF represents a simple oscillatory mode embedded in the signal. The decomposition is based on the local characteristic time scales of the data itself, rather than predefined basis functions like sinusoids or wavelets. The process involves an iterative ”sifting” procedure to extract the highest-frequency IMF

first, then subtracting it and repeating the process on the remainder until only a monotonic trend is left.

2. **Hilbert Spectral Analysis:** The Hilbert Transform is then applied to each IMF, $c_i(t)$, to form its analytic signal $z_i(t)$:

$$z_i(t) = c_i(t) + j\mathcal{H}\{c_i(t)\} = a_i(t)e^{j\theta_i(t)} \tag{C.3}$$

   where $\mathcal{H}\{\cdot\}$ denotes the Hilbert Transform. From the analytic signal, the instantaneous amplitude $a_i(t)$ and instantaneous phase $\theta_i(t)$ are obtained. The instantaneous frequency $\omega_i(t)$ for each IMF is then calculated as the time derivative of its phase:

$$\omega_i(t) = \frac{d\theta_i(t)}{dt} = \frac{d}{dt}\arg(z_i(t)) \tag{C.4}$$

The final HHT result is a time-frequency representation, often called the Hilbert spectrum, where the instantaneous frequency and amplitude of each IMF are plotted as a function of time. This representation is useful for analyzing complex micro-Doppler signatures for which fixed-resolution methods may be less effective.

Appendix D

Glossary of Key Technical Terms and Acronyms

This glossary provides definitions for key technical terms and acronyms used throughout this thesis, particularly in the context of radar systems, signal processing, and the coherent-on-receive methodology.

Term/Acronym	Definition
`ADC`	Analog-to-Digital Converter. An electronic device that converts a continuous analog signal (e.g., voltage) into a discrete sequence of digital numbers.
`ASCII`	American Standard Code for Information Interchange. A character encoding standard for electronic communication. Legacy Furuno files in this repository were sometimes stored as ASCII text, but the maintained workflow described in this thesis primarily uses binary IF captures.
`Autofocus`	In radar signal processing, particularly SAR and ISAR, autofocus refers to algorithms that estimate and correct residual phase errors to improve image sharpness or signal coherence. Earlier exploratory Furuno processing borrowed autofocus-style ideas, but the maintained pipeline in this thesis relies primarily on leakage-referenced phase correction, POC alignment diagnostics with pass-through timing, and CPI-wise phase tracking.
`AWGN`	Additive White Gaussian Noise. A basic noise model used in information theory to mimic the effect of many random processes that occur in nature. The "white" refers to a flat power spectral density, and "Gaussian" refers to a normal amplitude distribution.

Table D.1: Glossary of Key Technical Terms and Acronyms (Continued)

Term/Acronym	Definition
`Baseband`	In signal processing, baseband refers to the range of frequencies occupied by a signal before it is modulated to a higher carrier frequency (for transmission) or after it has been demodulated from a carrier frequency (at reception). For complex signals, it is centered around 0 Hz.
`Beat Frequency` (f_b)	In FMCW radar, the difference frequency generated by mixing the received signal with the currently transmitted signal (or a local oscillator tracking it). The beat frequency is proportional to the target's range (for stationary targets).
`BPF`	Bandpass Filter. A filter that passes frequencies within a certain range and rejects (attenuates) frequencies outside that range.
`CF`	Coherence Factor. A metric used to quantify the degree of phase stability or coherence across a set of signals (e.g., radar pulses). A value of 1 indicates perfect coherence, while 0 indicates complete incoherence. Defined as $\frac{1}{N_P}\lvert \sum_{p=1}^{N_P} e^{j\phi_p}\rvert$.
`CFO`	Carrier Frequency Offset. A difference between the actual carrier frequency of a received signal and the nominal or expected carrier frequency at the receiver, often due to oscillator inaccuracies or Doppler shifts. In this thesis, it primarily refers to pulse-to-pulse instabilities in the radar's effective carrier frequency or IF.
`Chirp`	A signal in which the frequency increases (up-chirp) or decreases (down-chirp) with time. Commonly used in FMCW radar and pulse compression.
`Clutter`	Unwanted radar echoes from objects that are not of interest, such as ground, sea, weather, or sometimes biological targets like birds when they are not the focus.
`Coherent Integration`	The process of summing multiple radar pulses while preserving their phase relationships. This improves the SNR for targets and enables Doppler processing.

Table D.1: Glossary of Key Technical Terms and Acronyms (Continued)

Term/Acronym	Definition
`Coherent-on-Receive`	A processing strategy in which an inherently non-coherent transmitter, such as a magnetron radar, is post-processed to recover phase-consistent data. In this thesis, coherence is induced by using the measured leakage signal as the physical phase reference within each pulse and, in the advanced extension, across CPIs.
`COTS`	Commercial-Off-The-Shelf. Refers to products that are readily available for purchase and not custom-designed.
`CPI`	Coherent Processing Interval. The total time duration over which a sequence of M radar pulses is coherently processed, typically to extract Doppler information. $T_{\text{CPI}} = M \cdot T_{PRI}$.
`CVD`	Cadence-Velocity Diagram. A secondary time-frequency representation that highlights periodic micro-motion by plotting modulation cadence against Doppler velocity.
`Cross-Correlation`	A measure of similarity between two signals as a function of relative displacement or lag. In the maintained Furuno workflow, global CPI-to-CPI phase comparison can be written as a Frobenius inner product across successive range-Doppler maps, while phase-only correlation is retained as a fast-time alignment diagnostic.
`Digital Downconversion`	The process of converting a digitized IF signal to a complex baseband signal using digital signal processing techniques, typically involving multiplication by complex sinusoids and digital filtering.
`Doppler Effect / Shift (`f_D`)`	The change in frequency of a wave in relation to an observer who is moving relative to the wave source. In radar, it's caused by the radial velocity of a target. $f_D = 2v_r/\lambda$.
`DSP`	Digital Signal Processing. The use of digital processing, such as by computers or more specialized digital signal processors, to perform a wide variety of signal processing operations.

Table D.1: Glossary of Key Technical Terms and Acronyms (Continued)

Term/Acronym	Definition
`Fast-Time`	In pulsed radar data, fast-time refers to the samples collected within a single pulse repetition interval (PRI), corresponding to echoes from different ranges. The sampling rate is the ADC sampling rate.
`FFT`	Fast Fourier Transform. An efficient algorithm to compute the Discrete Fourier Transform (DFT) and its inverse.
`FMCW`	Frequency Modulated Continuous Wave. A type of radar that transmits a continuous signal whose frequency is modulated (typically linearly, as a chirp).
`FSST`	Fourier Synchrosqueezed Transform. A reassigned time-frequency representation that sharpens the concentration of chirp-like or oscillatory components relative to a standard STFT.
`Furuno FAR-15x8`	A specific series of commercial X-band marine navigation radar systems manufactured by Furuno, used as the data source in this thesis.
`IF`	Intermediate Frequency. In a superheterodyne receiver, the RF signal is converted to a lower, fixed frequency (the IF) for easier amplification and filtering before demodulation. For the Furuno radar in this work, the nominal IF is 60 MHz.
`IIR Filter`	Infinite Impulse Response Filter. A type of digital filter whose impulse response is non-zero over an infinite length of time. Butterworth filters are an example.
`I/Q Data`	In-phase (I) and Quadrature (Q) components of a complex signal. These represent the real and imaginary parts of the complex baseband signal.
`Leakage (Transmit Leakage)`	A portion of the transmitted radar pulse energy that directly couples into the receiver path without reflecting off a target, typically due to imperfect isolation in the TR switch or circulator. It appears at or near zero range.
`LPF`	Low-Pass Filter. A filter that passes signals with a frequency lower than a selected cutoff frequency and attenuates signals with frequencies higher than the cutoff frequency.

Table D.1: Glossary of Key Technical Terms and Acronyms (Continued)

Term/Acronym	**Definition**
`Magnetron`	A high-powered vacuum tube that generates microwaves using the interaction of a stream of electrons with a magnetic field while moving past a series of open metal cavities (cavity resonators). Used as the transmitter source in many non-coherent radars.
`Matched Filter`	An optimal linear filter used in radar to maximize the signal-to-noise ratio (SNR) for a known signal in the presence of additive stochastic noise. Its impulse response is a time-reversed, conjugated version of the input signal.
`MAD`	Median Absolute Deviation. A robust measure of the variability of a univariate sample of quantitative data. It is defined as the median of the absolute deviations from the data's median.
`Micro-Doppler Effect`	Frequency modulations superimposed on the main Doppler shift of a target, caused by mechanical vibrations or rotations of parts of the target (e.g., helicopter rotor blades, bird wings, human limbs).
`MTI`	Moving Target Indication. A radar technique used to reject clutter (stationary echoes) and detect moving targets by exploiting their Doppler shift. Traditional MTI often relies on inter-pulse phase cancellation.
`Non-Coherent Radar`	A radar system in which the phase of the transmitted signal is not precisely controlled or known from pulse to pulse, typically due to the use of a free-running oscillator like a magnetron.
`PRF` (f_{PRF})	Pulse Repetition Frequency. The number of radar pulses transmitted per second. $f_{\mathrm{PRF}} = 1/T_{PRI}$.
`PRI` (T_{PRI})	Pulse Repetition Interval. The time duration between the start of one radar pulse and the start of the next consecutive pulse.
`Pulse Compression`	A signal processing technique used in radar to increase the range resolution of long pulses. It involves transmitting a long pulse with internal modulation (e.g., LFM or phase coding) and then processing the received echo with a matched filter to compress the energy into a short effective pulse.

Table D.1: Glossary of Key Technical Terms and Acronyms (Continued)

Term/Acronym	**Definition**
`POC`	Phase-Only Correlation. A registration method that estimates relative timing or shift from the phase of the cross-spectrum rather than from raw amplitude correlation. In this thesis it is retained as a diagnostic and future guarded alignment option; the reported maintained runs use pass-through timing.
`Quadratic Interpolation`	A method of interpolation where the interpolant is a quadratic polynomial. Used in this thesis to refine the IF estimate from an FFT spectrum.
`Range Bin`	A discrete cell in range, corresponding to a specific time delay interval after pulse transmission, from which radar echoes are sampled. The width of a range bin is ideally matched to the radar's range resolution.
`Range-Doppler Map (RDM)`	A two-dimensional representation of radar data, where one axis represents target range and the other represents Doppler frequency (or velocity). Target echoes appear as peaks in this map.
`RCS` (σ)	Radar Cross Section. A measure of how detectable an object is by radar. It is a fictitious area that represents the intensity of the signal scattered back to the radar from the target.
`RMSE`	Root Mean Square Error. A frequently used measure of the differences between values (sample or population values) predicted by a model or an estimator and the values observed. In this thesis, used to quantify residual phase errors.
`Slow-Time`	In pulsed radar data, slow-time refers to the dimension across multiple pulses for a fixed range bin. The sampling rate in slow-time is the PRF. Doppler information is extracted from slow-time samples.
`SNR`	Signal-to-Noise Ratio. The ratio of the power of a signal (meaningful information) to the power of background noise.
`SOS`	Second-Order Sections. A way of implementing higher-order digital IIR filters as a cascade of biquadratic (second-order) filter sections, which improves numerical stability.

Table D.1: Glossary of Key Technical Terms and Acronyms (Continued)

Term/Acronym	Definition
`Spectrogram`	A visual representation of the spectrum of frequencies of a signal as it varies with time. Generated using the Short-Time Fourier Transform (STFT).
`STFT`	Short-Time Fourier Transform. A Fourier-related transform used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time.
`SVD`	Singular Value Decomposition. A factorization of a real or complex matrix that generalizes the eigendecomposition of a square normal matrix to any $m \times n$ matrix. In this thesis it is used for leakage-template extraction, clutter-aware shrinkage, and spectrogram denoising.
`Trigger (TRIG_AT)`	A timing reference signal in the data acquisition system that indicates the start of a radar pulse transmission or a related event, used for segmenting the continuous data stream.
`Unambiguous Range` (R_{ua})	The maximum range to which a transmitted radar pulse can travel and return to the radar before the next pulse is transmitted. $R_{ua} = cT_{PRI}/2$.
`Unambiguous Velocity` (v_{ua})	The maximum radial velocity that a radar can measure without aliasing. $v_{ua} = \pm\lambda f_{\mathrm{PRF}}/4$.
`Unwrap (Phase)`	The process of reconstructing the original continuous phase angle from its principal values (which are typically wrapped to the interval $[-\pi, \pi]$ or $[0, 2\pi]$) by adding appropriate multiples of 2π.
`X-band`	A segment of the microwave radio region of the electromagnetic spectrum, typically defined as frequencies from 8.0 to 12.0 GHz.