Machine Learning-based Automatic Modulation Classification for 5G-Advanced and 6G Waveforms: Robust Identification Under Realistic Channel Impairments

Abdelkader Horch, Mokhtar Besseghier, Samir Ghouali, and Abderrahmane Louni

University of Mustapha Stambouli, Mascara, Algeria

https://doi.org/10.26636/jtit.2026.3.2595

Abstract — Automatic modulation classification (AMC) for 5G-Advanced and 6G networks must blindly identify waveforms from received signals under realistic channel impairments, enabling cognitive radio dynamic spectrum access and interference avoidance. No prior work has simultaneously applied machine learning to classify all eight leading waveforms (UFMC, GFDM, FBMC, NOMA, OFDM-IM, OTFS, ODDM, and AFDM) under realistic channel impairments, nor quantified the minimum feature set for resource-constrained deployment. We present a framework that (i) extracts a 38-dimensional feature vector that includes three novel channel-aware characteristics (amplitude fading variance, phase discontinuity, and frequency drift); (ii) benchmarks nine machine learning classifiers, including an FC-MLP deep learning baseline and five feature selection methods, on 201 600 signals across twelve channel conditions (nine custom plus three 3GPP TDL profiles) and seven SNR levels, with leakage-free feature selection; and (iii) identifies a compact 10-feature subset validated with Bonferroni-corrected McNemar tests and Wilson confidence intervals. FC-MLP achieves 99.09% accuracy; ensemble-boost (99.04%) and random forest (99.02%) are statistically equivalent. The 10-feature random forest reaches a score of 98.90% within 0.12 pp of the full feature baseline at a cost that is 74% lower and with a 0.071 ms inference per block. The five-fold cross-validation confirms stability (98.54%, Wilson 95% CI: 98.49%, 98.59%). Per channel accuracy ranges from 98.87% (Rayleigh) to 99.98% (AWGN/Rician); 3GPP TDL-A/B/C profiles confirm transferability to 5G NR. The three channel-aware features yield up to 3.1% gain under double-selective fading and an average overall improvement.

Keywords* — *5G-Advanced/6G, automatic modulation classification, cognitive radio, compact feature selection, dynamic spectrum access, machine learning, realistic channel impairments

1. Introduction

Automatic modulation classification (AMC) enables a receiver to blindly determine the waveform type of an observed signal without prior knowledge of the modulation scheme, channel conditions, or signal parameters [1], [2]. In the context of 5G-Advanced and 6G networks, robust AMC is indispensable for a wide range of applications, including access to the dynamic spectrum of cognitive radio (CR), interference avoidance, cooperative co-existence and spectrum policy enforcement [3]–[6]. Despite its importance, AMC for the full set of next-generation waveform candidates under realistic channel propagation conditions remains a largely open problem.

The 5G-Advanced and 6G standardization processes have produced a diverse set of waveform candidates, each targeting specific deployment requirements: UFMC and GFDM for flexible subband allocation, FBMC for high spectral efficiency without cyclic prefix, NOMA for massive connectivity via power domain multiplexing; OFDM-IM for energy-efficient index modulation; and OTFS, ODDM, and AFDM for robust high-mobility communications via delay-Doppler or chirp-based signal representations [7]–[10]. Precise classification of these waveforms is present, using machine learning on blind feature observations, and is a prerequisite for all subsequent spectrum management decisions. Recent AMC studies have extensively addressed legacy modulation schemes such as PSK, QAM, and FSK, but have not simultaneously tackled all eight next-generation waveform candidates under compound realistic channel impairments.

1.1. Motivation and Challenges

The deployment of ML-based AMC for next-generation waveforms faces three distinct difficulties: structural similarity among the waveform candidates (which challenges any classifier), severe distortion from realistic propagation channels (which corrupts the most informative features), and strict computational constraints when AMC must operate in real time, which limits model complexity. These challenges are especially acute in cognitive radio and spectrum monitoring scenarios, where the receiver operates independently from the transmitter.

Structural similarity. Several next-generation waveforms share overlapping time-frequency characteristics, making blind ML classification particularly difficult. OTFS, ODDM, and AFDM are all designed for Doppler resilience through related delay-Doppler or chirp-based representations, making their spectral signatures difficult to distinguish. FBMC, GFDM, and UFMC all exploit subband filtering, mainly differing in prototype filter design. Previous AMC methods trained under clean channel conditions suffer 5 – 15% accuracy degradation when evaluated under realistic impairments [1], [10], as structurally discriminative characteristics become corrupted by channel effects, i.e., precisely the conditions any real AMC system must handle.

Realistic channel impairments. A CR receiver cannot control or predict the channel it observes. The waveform identification module must therefore be robust enough to face the following:

- Multipath fading: Rayleigh and Rician fading introduce time-varying amplitudes and phase distortions that obscure waveform-specific envelope statistics.
- Doppler effects: high mobility CR scenarios (30 – 200 km/h) induce frequency dispersion and intercarrier interference, particularly harmful for subcarrier-structured waveforms.
- Frequency-selective channels: delay spread exceeding the symbol period causes inter-symbol interference that disrupts cyclic prefix and periodicity features.
- Synchronization errors: a CR receiver operating without cooperation from the transmitter cannot assume perfect carrier or timing synchronization; residual CFO and timing misalignment distort the phase and cyclo-stationary structures of all waveforms.
- Combined impairments: simultaneous time-frequency double-selective channels – common in vehicular and UAV CR deployments – present the most challenging scenario.

Selection of compact features for real-time AMC. A key open question in ML-based AMC is: what is the minimum feature set that preserves near-maximum accuracy under a rigorous, leakage-free selection protocol? This work attempts to answer this question for any resource-constrained deployment, including CR platforms such as IoT gateways, UAV-mounted spectrum sensors, and battery-powered secondary nodes, with all of them imposing strict limits on resources and latency. A waveform classifier that requires hundreds of features or seconds of inference latency is not suitable for real-time operation. No prior work has rigorously quantified the minimum feature set needed for next-generation waveform AMC under realistic channels, nor validated that such a compact model can operate within real-time latency budgets.

The above considerations lead to the following research questions:

1) Which machine learning algorithm achieves the best accuracy-complexity trade-off for AMC of eight 5G-Advanced/6G waveforms under realistic channel conditions, satisfying real-time latency requirements?
2) What is the minimum feature subset that preserves near-maximum AMC accuracy under a leakage-free selection protocol? Does it generalize to ML classifier architectures?
3) Do channel-aware features explicitly designed to capture propagation impairments provide statistically significant AMC improvements, as confirmed by McNemar tests?

1.2. Research Contribution

This paper proposes a CR-oriented 38-feature blind identification framework with mathematical definitions for all features, operating in a fully blind mode, as required by CR receivers with no transmitter cooperation.

We introduce three new channels-aware features (F36 – F38): amplitude fading variance, phase discontinuity, and frequency drift, explicitly designed to characterize the uncontrolled propagation environment a CR receiver observes. Ablation studies quantify their contribution: removing all three degrades identification accuracy by 1.5% on average and by up to 3.1% under double selective fading.

We benchmark nine classifiers, including an FC-MLP deep learning baseline and five feature-selection methods on 201 600 signals across twelve channel conditions and seven SNR levels. FC-MLP achieves the highest accuracy at 99.09%; random forest provides the best accuracy-per-feature efficiency for CR deployment. McNemar tests (corrected with Bonferroni) confirm the statistical significance of all reported differences.

The work extends evaluation to 3GPP standard channels. Three TDL profiles (TDL-A, TDL-B, TDL-C) of 3GPP TR 38.901 [11] are added along with nine custom channel conditions, enabling direct comparability with 5G NR deployment scenarios that CR devices will encounter. All feature importance scoring is performed strictly within the training fold; the test set plays no role in feature ranking, ensuring that the reported 10-feature accuracies are genuine out-of-sample estimates suitable for real CR deployment decisions.

The results demonstrate a 74% computational drop by reducing from 38 to 10 features with an accuracy loss of only 0.12% (McNemar $p > 0.05$), achieving 0.071 ms per segment inference and meeting the real-time latency requirements of CR dynamic spectrum access. The compact model is transferred to all evaluated classifiers. We validate robustness through parametric analysis covering SNR (5 to +25 dB), Doppler (0 – 120 km/h), CFO (0 – 5 kHz) and timing offset (0 – 15 samples), i.e., the full envelope of impairments a CR receiver must tolerate, with per waveform confusion analysis identifying the most challenging identification pairs.

The paper is organized as follows. Section 2 describes the system model and the 12-channel CR evaluation setup. Section 3 details the blind feature extraction framework. Section 4 presents classification results including the FCMLP baseline test and the McNemar test. Section 5 analyzes the importance of leak-free features. Section 6 evaluates the reduction for CR deployment. Section 7 reports on robustness analysis. Section 8 discusses the implications for the deployment of CR and Section 9 concludes the study.

2. System Model and Dataset Generation

2.1. Signal Model

We model a CR receiver that passively observes a received signal block and is required to identify the waveform without any cooperation from the transmitter. Consider a transmitted signal $s[n]$ corresponding to one of the eight 5G-Advanced/6G waveform candidates, sampled at frequency f_s. The complex baseband signal at the CR receiver after channel propagation is:

$$r[n] = \sum_{\ell=0}^{L-1} h_\ell[n]\, s[n-\ell] + w[n] \ , \quad (1)$$

where $h_\ell[n]$ is the time-varying complex gain of the ℓ-th multipath component and $w[n] \sim \mathcal{CN}(0, \sigma_w^2)$ is complex AWGN with

$$\sigma_w^2 = P_s/\text{SNR} \ ,$$
$$P_s = \mathbb{E}[|s[n]|^2] \ .$$

All twelve channel configurations are specific instantiations of Eq. (1).

2.2. Channel Models

AWGN: $h_0[n] = 1$ and all other paths zero. Baseline line-of-sight scenario.

Rayleigh fading: Non-line-of-sight single-path propagation with Doppler-correlated complex gain:

$$h[n] = \frac{1}{\sqrt{2}}\big(g_I[n] + jg_Q[n]\big) \ , \quad (2)$$

where $g_I[n]$ and $g_Q[n]$ are correlated Gaussian processes whose power spectral density follows the Clarke/Jakes model

$$S(f) = [\pi f_d \sqrt{1-(f/f_d)^2}]^{-1} \text{for} |f| \leqslant f_d \ ,$$

with temporal autocorrelation

$$R_h[m] = J_0(2\pi f_d\, mT_s) \ ,$$

where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind and $T_s = 1/f_s$.

The 16-oscillator Jakes simulator [12] is used for implementation. For the static Rayleigh channel $f_d \to 0$ yields independent identically distributed realizations.

Rician fading: Line-of-sight with K factor of 10 dB incorporating the LOS Doppler shift:

$$h[n] = \sqrt{\frac{K}{K+1}}\, e^{j2\pi f_{\text{LOS}} n T_s} + \sqrt{\frac{1}{K+1}}\, h_R[n] \ , \quad (3)$$

where $f_{\text{LOS}} = f_d \cos\theta$ is the LOS Doppler frequency ($\theta = 0°$ for worst-case maximum Doppler) and $h_R[n]$ is a scattered component following the Rayleigh model.

Frequency offset: Post-synchronization residual CFO of $\Delta f = 150$ Hz (corresponding to ≈ 0.063 ppm at $f_c =$ 2.4 GHz) is applied as

$$r_{\text{CFO}}[n] = s[n]\, e^{j2\pi \Delta f n T_s} .$$

This models the residual offset after initial carrier synchronization. Pre-synchronization scenarios with larger offsets (up to 5 kHz, i.e., ≈ 2 ppm) are evaluated in Sec. 7.

Timing offset: Integer sample timing misalignment of $\tau_0 = 5$ samples $r_\tau[n] = r[n-\tau_0]$.

Low Doppler: Jakes model [12] with $v = 30$ km/h, $f_d \approx$ 67 Hz, urban mobile scenario.

High Doppler: Jakes model with $v = 120$ km/h, $f_d \approx$ 267 Hz, high-speed vehicular scenario.

Frequency-selective: $L = 6$ paths with exponential power delay profile $\mathbb{E}[|h_\ell|^2] \propto e^{-\ell/2}$ (unit total power), RMS delay spread $\tau_{\text{rms}} = 300$ ns, i.e. 2.4 samples at 8 kHz, within the CP length $N_{\text{CP}} = 16$ samples of CP-based waveforms.

Time-frequency double-selective: Combined Doppler ($v =$ 60 km/h, $f_d \approx 133$ Hz) and frequency-selective multipath ($\tau_{\text{rms}} = 100$ ns, $L = 4$ paths).

3GPP TDL Standard Profiles (TDL-A, TDL-B, TDL-C): To ensure the CR waveform identifier transfers directly to real 5G NR deployment scenarios [11], three standardized tapped delay line profiles from 3GPP TR 38.901 are included. Each profile applies an independent complex Gaussian gain per tap (Rayleigh per tap), with delays and powers taken verbatim from the standard:

- TDL-A (NLOS, $\tau_{DS} = 30$ ns): 23 taps that model indoor hotspots and dense urban propagation, typical of small cell or IoT gateway deployments,
- TDL-B (NLOS, $\tau_{DS} = 100$ ns): 22 taps that model urban macro propagation, representative of vehicular CR and mobile spectrum monitoring scenarios,
- TDL-C (NLOS, $\tau_{DS} = 300$ ns): 24 taps that model rural macro and large delay spread scenarios, the most severe frequency selective profile and most challenging for the identification of CR.

The generalized TDL channel is implemented as:

$$r[n] = \sum_{p=1}^{P} g_p\, s[n-\tau_p] + w[n] \ , \quad (4)$$

where τ_p is the p-th tap delay (in samples, rounded from the nanosecond values in the standard), $g_p \sim \mathcal{CN}(0, \sigma_p^2)$ with $\sigma_p^2 \propto 10^{\Delta_p/10}$ and Δ_p the relative power in dB, normalized so $\sum_p \sigma_p^2 = 1$. The output power is rescaled to match input power after convolution.

2.3. Waveform Configurations and Signal Models

All eight waveforms use 4-QAM symbols. The received block $r[n]$, $n = 0, \ldots, N-1$, is the sole input to feature extraction, reflecting the CR receiver's blind observation. The identifier has no access to waveform type, channel parameters, or SNR as mandated by the CR non-cooperative sensing paradigm.

The universal filtered multi-carrier (UFMC) signal is defined as:

$$s_{\text{UFMC}}[n] = \sum_{b=1}^{B} \sum_{m=0}^{M_b-1} X_b[m]\, g_b[n-mL]\, e^{j2\pi f_b n}\,, \tag{5}$$

$B = 8$ subbands, $M_b = 8$ subcarriers per subband, $g_b[n]$ is a 43-tap Dolph-Chebyshev FIR filter (-40 dB sidelobe suppression), $L = 1$.

Subband center frequencies

$$f_b = (b-1)\Delta f_{SB}$$

with subband spacing

$$\Delta f_{SB} = \frac{M_b f_s}{K} = 1\ \text{kHz}$$

$K = 64$ total subcarriers, $\Delta f_{sc} = f_s/K = 125$ Hz. The block length is 256 samples.

The generalized frequency division multiplexing signal is:

$$s_{\text{GFDM}}[n] = \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} d_{k,m}\, g\big[(n-mK) \bmod N\big]\, e^{j2\pi \frac{k}{K} n}\,, \tag{6}$$

where $K = 64$ subcarriers, $M = 8$ time slots, raised-cosine pulse with roll-off $\alpha = 0.1$, $N = KM = 512$ samples [13], [14].

The filter bank multi-carrier with OQAM (FBMC-OQAM) signal is defined as:

$$s_{\text{FBMC}}[n] = \sum_{k=0}^{K-1} \sum_{m\in\mathbb{Z}} a_{k,m}\, p[n-mK]\, e^{j\frac{\pi}{2}k}\, e^{j2\pi \frac{k}{K} n}\,, \tag{7}$$

where $K = 64$, $a_{k,m} \in \mathbb{R}$ real-valued OQAM symbols, the prototype PHYDYAS filter of length $4K = 256$, stagger offset $K/2$ [15], [16]. The block length is 512 samples.

The power-domain superposition non-orthogonal multiple access (NOMA) signal is defined as follows:

$$s_{\text{NOMA}}[n] = \sum_{k=0}^{K-1} \big(\sqrt{\alpha_1} X_1[k] + \sqrt{\alpha_2} X_2[k]\big) e^{j2\pi \frac{k}{K} n}\,, \tag{8}$$

where $K = 256$, power allocation $\alpha_1 = 0.8, \alpha_2 = 0.2$ [17]–[19]. The block length is 256 samples.

The OFDM with index modulation (OFDM-IM) signal is the following:

$$s_{\text{OFDM-IM}}[n] = \frac{1}{\sqrt{K}} \sum_{g=1}^{G} \sum_{k\in\mathcal{I}_g} X_g[k]\, e^{j2\pi \frac{k}{K} n}\,, \tag{9}$$

where $K = 64$ total subcarriers, $G = 12$ groups of 5 subcarriers (4 guard subcarriers), $K_a = 3$ active subcarriers per group (60% activation).

The active index set $\mathcal{I}_g$ is selected by $\lfloor \log_2 \binom{5}{3} \rfloor = 3$ index bits via a combinatorial number system lookup table [20]–[22]. A 16-sample cyclic prefix is appended. The block length is 80 samples.

The orthogonal time-frequency space (OTFS) modulates in the Doppler delay domain [8], [23], [24]. The 2D ISFFT grid signal is defined as follows:

$$x[n,m] = \frac{1}{\sqrt{MN}} \sum_{k=0}^{N-1} \sum_{\ell=0}^{M-1} X[k,\ell]\, e^{j2\pi\left(\frac{nk}{N} - \frac{m\ell}{M}\right)}\,, \tag{10}$$

with $N = M = 8$ (8×8 grid).

The 2D array is serialized to a 1D sequence via the Heisenberg transform:

$$s_{\text{OTFS}}[i] = x[i \bmod N, \lfloor i/N \rfloor], \quad i = 0, \ldots, 63.$$

The block length is 72 samples.

The orthogonal delay Doppler multiplexing (ODDM) signal is the following:

$$s_{\text{ODDM}}[n] = \frac{1}{\sqrt{K}} \sum_{k=0}^{K-1} \Big(\sum_{s=0}^{S-1} H_{k,s}\, d_s \Big) e^{j2\pi \frac{k}{K} n}\,, \tag{11}$$

where $K = 64$, spreading factor $S = 4$, Walsh-Hadamard matrix $\mathbf{H} \in \mathbb{C}^{K\times S}$, 16-sample CP [25]. The block length is 80 samples.

The chirp-based affine frequency division multiplexing (AFDM) signal is defined as:

$$s_{\text{AFDM}}[n] = \sum_{k=0}^{K-1} X[k]\, e^{j\pi\mu n^2}\, e^{j2\pi \frac{k}{K} n}\,, \tag{12}$$

where $K = 128$ chirp subcarriers, optimized chirp rate μ for Doppler tolerance $f_d \leqslant 500$ Hz Doppler tolerance [26]–[28]. The block length is 128 samples.

The sampling frequency is $f_s = 8$ kHz. For multicarrier waveforms, the relevant timing parameter is the symbol duration (e.g., $T_{\text{sym}} = K/f_s = 8$ ms for $K = 64$ subcarriers). Per-waveform block lengths are shown in Tab. 1. The 8 kHz rate is a reduced-scale proof of concept that preserves all waveform structural properties. Scalability to 5G wideband 5G bandwidths is discussed in Subsection 8.3.

2.4. Dataset Statistics

The dataset comprises 8 waveform types × 7 SNR levels $[-5, 0, 5, 10, 15, 20, 25]$ dB × 12 channel conditions × 300 realizations = 201 600 signals in total, balanced at 25 200 per waveform class. The original 9-channel subset of 151 200 signals is retained for direct comparison with previous work.

The signals use independent random seeds derived from master seed 42. A stratified 70 / 30% train-test split maintains SNR and channel diversity in both partitions. The identical partition object is reused across all pipeline steps to prevent any cross-contamination.

3. Feature Extraction Framework

Used hereafter, the term *feature* refers exclusively to a numerical descriptor computed from the received block $r[n]$ and used as a machine learning input.

The framework operates on the received block $r[n]$, $n = 0, \ldots, N-1$, in the fully blind mode. The 38 feature definitions are provided below for reproducibility.

Tab. 1. Simulation parameters used to analyze the proposed system.

Waveform	Eq.	Subcarriers	CP/guard	Block	Characteristic[†]
UFMC	(5)	64 (8 bands)	43-tap filter	256	Per-subband FIR filtering
GFDM	(6)	64	8 slots	512	Circular pulse shaping
FBMC	(7)	64	4× ovlp	512	OQAM, prototype filter
NOMA	(8)	256	—	256	Power domain superposition
OFDM-IM	(9)	64	16 samples	80	Index modulation (60%)
OTFS	(10)	8×8	8 samples	72	Delay-Doppler domain
ODDM	(11)	64	16 samples	80	Hadamard spreading
AFDM	(12)	128	—	128	Chirp-based, Doppler-robust

[†] Waveform characteristic describes the signal-processing property of each waveform design. It is not a machine learning input feature (see Sec. 3).

3.1. Classical Signal Features (F1 – F35)

Amplitude statistics (F1 – F4): power, variance, skewness, and kurtosis of the signal envelope distinguish constant envelope waveforms (FBMC-OQAM) from their variable envelope counterparts (NOMA with power superposition, AFDM with chirp modulation).

Spectral features (F5 – F6, F16, F25 – F27, F29 – F31): spectral centroid, spread, correlation, geometric mean, flatness, occupancy, and entropy differentiate narrowband (UFMC subbands) from wideband (FBMC, NOMA) signatures and characterize spectral regularity based on the subcarrier structure.

Instantaneous frequency features (F7 – F11, F14): statistics of $\phi'[n] = \Delta(\text{unwrap}(\arg(r[n])))$ capture frequency modulation characteristics. Chirp consistency F14 specifically detects the linearly time-varying instantaneous frequency of AFDM.

Algorithm 1 Robust AMC framework (leakage-free)

Require: Received blocks $\{r_i[n]\}_{i=1}^{N_s}$, labels $\{y_i\}$
Ensure: Trained classifier $\hat{f}(\cdot)$, optimal feature indices $\mathcal{S}^*$

1: **Dataset generation:** Simulate per Tab. 1 and Sec. 2
2: **Feature extraction:** For each block $r_i[n]$
3: Compute DFT: $R[k] = \text{DFT}(r[n])$; instantaneous phase $\phi[n] = \text{unwrap}(\arg(r[n]))$
4: Compute instantaneous frequency $\phi'[n] = \Delta\phi[n]$
5: Compute time-frequency representations (STFT, WVD, FrFT)
6: Extract f_1–f_{35} (Tab. 2) and f_{36}–f_{38}, Eqs. (13) – (15)
7: **Train/test split:** Stratified 70 / 30% partition $\mathcal{P}$; save $\mathcal{P}$ for reuse
8: **Normalization:** $\tilde{f}_i = (f_i - \mu_{\text{train}})/\sigma_{\text{train}}$
9: **Feature selection (leakage-free):** Apply Alg. 2 to training fold only; select top-K features $\mathcal{S}^*$
10: **Training:** Train all classifiers on $\tilde{f}_{i,\mathcal{S}^*}$ using same $\mathcal{P}$; hyperparameters via inner 5-fold CV on training fold
11: **Evaluation:** Accuracy, precision, recall, F1, McNemar tests, Wilson CI on 30% test set
12: **Return** $\hat{f}(\cdot)$, $\mathcal{S}^*$

Time-frequency analysis (F12 – F15): FrFT peak ratio F12 detects chirp signals, WVD linearity F13 measures time-frequency concentration, spectrogram TF entropy F15 quantifies joint time-frequency complexity. Note: WVD calculation $\mathcal{O}(N^2 \log N)$ is the costliest feature to compute. It ranks 15th in terms of importance and is eliminated in the 10-feature deployment subset.

Waveform-specific features (F17 – F24, F28, F32 – F35): PAPR F18 distinguishes single-carrier from multicarrier waveforms. CP correlation F24 identifies CP-OFDM structures. Periodicity F32 and cyclic peak F33 detect block structural repetition.

3.2. New Channel-aware Features (F36 – F38)

Three features explicitly designed to capture propagation effects, invariant to waveform structure:

Amplitude fading variance (F36):

$$F_{36} = \frac{\text{Var}(P_{\text{local}})}{\big(\mathbb{E}[P_{\text{local}}]\big)^2}\,, \tag{13}$$

where:

$$P_{\text{local}}[k] = \frac{1}{W}\sum_{n=kW}^{(k+1)W-1} \big|r[n]\big|^2\,,$$

$$W = \min\big(32, \lfloor N/4 \rfloor\big)\,.$$

It is near zero under AWGN ($F_{36} < 0.1$) and elevated under Rayleigh/Rician fading ($F_{36} > 0.3$).

Phase discontinuity (F37):

$$F_{37} = \text{std}\big(\phi''[n]\big)\,, \tag{14}$$

where

$$\phi''[n] = \Delta^2\big(\text{unwrap}(\arg(r[n]))\big)$$

is the second order phase difference. Pearson correlation with CFO magnitude: $\rho = 0.82$.

Frequency drift (F38):

$$F_{38} = \text{std}(f_{\text{local}})\,, \tag{15}$$

where

$$f_{\text{local}}[k] = \frac{1}{W} \sum_{n=kW}^{(k+1)W-1} \phi'[n]$$

increases 3 – 5× from 30 km/h to 120 km/h Doppler scenarios.

Ablation studies (Tab. 6 in Subsection 5.2) confirm that the removal of F36 – F38 degrades accuracy by up to 3.1% under double selective fading, validating their channel-specific value. F36 ranks 7th out of 38 features (ensemble score 0.360); F37 ranks 21st (score 0.190). F38 ranks 36th out of 38 (score 0.103) and does not enter the 10-feature deployment subset. Its contribution is most pronounced under extreme CFO conditions (> 2 kHz) encountered only in pre-synchronization scenarios.

4. Multi-algorithm Classification Study

4.1. Evaluated Classifiers

Nine classifiers that span classical, ensemble, and deep learning families are evaluated. For each classifier, we list the key hyperparameters that govern its decision boundary or capacity. The full grid-search details are given in Sec. 4.2:

1) k-NN: $k = 5$, Euclidean distance; majority vote on five nearest neighbors. Simple non-parametric baseline.
2) SVM-linear: ECOC one-versus-all, linear kernel, regularization $C = 1$. Decision boundaries are hyperplanes in the 38-D feature space.
3) SVM-RBF: Gaussian kernel (RBF), $C = 10$, automatic kernel scale estimation ($\gamma =$ auto). Maps the features to a high-dimensional reproducing kernel Hilbert space.
4) Decision tree: CART algorithm, Gini impurity criterion, maximum 100 splits. Single-tree baseline; the shallowest interpretable model evaluated.
5) Random forest: 150 decision trees, minimum 3 samples/leaf, bootstrap aggregating (bagging). Provides feature importance scores used in Sec. 5.
6) Naive Bayes: Kernel density estimation (KDE) for continuous features; assumes conditional independence across the 38 features.
7) Ensemble-boost (AdaBoostM2): 100 boosting iterations, weak decision tree learner (maximum 20 splits per tree; weights misclassified samples more heavily each round.
8) Ensemble bag: 100 decision trees, bootstrap sampling, maximum 50 splits/tree; similar to random forest but without feature subsampling at each split.
9) FC-MLP (deep-learning baseline): Fully-connected network $256 \to 128 \to 64 \to 8$ [29]; batch normalization, ReLU, dropout $p = 0.3$ per hidden layer; Adam optimizer, ℓ_2 regularization $\lambda = 10^{-4}$, 100 epochs. Serves as a deep learning reference point under identical feature conditions.

All classifiers use z-score normalized features

$$\tilde{f} = \frac{f - \mu_{\text{train}}}{\sigma_{\text{train}}},$$

training-set statistics only. Random seed 42 for reproducibility.

4.2. Hyperparameter Optimization

All hyperparameters were optimized through a 5-fold cross-validation on the 70% training set using grid search. Key results: optimal k-NN $k = 5$ from $\{1, 3, 5, 7, 10, 15\}$; SVM-linear $C = 1$ from $\{0.01, 0.1, 1, 10, 100\}$; SVM-RBF $C = 10$, $\gamma =$ auto; decision tree maximum 100 splits; random forest 150 trees, min leaf 3; ensemble-boost 100 iterations, max 20 splits; ensemble-bag 100 trees, max 50 splits.

FC-MLP: initial learning rate from $\{10^{-4}, 10^{-3}, 10^{-2}\}$ (optimal 10^{-3}); dropout from $\{0.2, 0.3, 0.5\}$ (optimal 0.3); ℓ_2 weight decay from $\{10^{-5}, 10^{-4}, 10^{-3}\}$ (optimal 10^{-4}); architecture depth 2 vs. 3 hidden layers and mini-batch size $\{128, 256\}$ also evaluated. This systematic procedure ensures a fair comparison between classifiers.

4.3. Performance Metrics

We report accuracy, i.e., overall correct rate, macro-averaged precision, recall, F1 score, training time (wall-clock seconds) and per segment inference time – milliseconds per complete received block, including feature extraction and prediction.

4.4. Experimental Results

Key observations are as follows:

FC-MLP achieves the highest accuracy of 99.09%, marginally outperforming ensemble boost (99.04%) and random forest (99.02%). McNemar tests show that the FC-MLP vs. random forest gap ($p = 0.039$) and FC-MLP vs. ensemble-boost gap ($p = 0.082$) are both not statistically significant after Bonferroni correction $\alpha_B = 0.05/36 \approx 0.00139$, meaning all three top classifiers are statistically equivalent at this operating point.

Ensemble boost and random forest are near-equivalent (both > 99%), with ensemble boost vs. random forest gap ($p = 0.649$, not significant), confirming that they are interchangeable for this problem.

SVM-RBF requires the longest training time (2 545.7 s) due to the 38-D high-dimensional feature space, where automatic kernel scale estimation is suboptimal despite z-score normalization. Naive Bayes has the highest inference latency of 51.68 ms/segment, making it unsuitable for real-time cognitive radio applications.

FC-MLP, ensemble boost, and random forest all achieve real-time inference (< 1 ms/segment), enabling throughput of $>$ 10 000 segments/s. Note that the 0.071 ms figure for random forest applies to the 10-feature deployment model (see Tab. 3 footnote); the 38-feature RF baseline achieves 0.021 ms inference.

Decision tree is the best lightweight option (98.70%, 3.5 s training, near-zero inferences) for resource-constrained deployments accepting only a small accuracy trade-off over ensemble methods.

Tab. 2. Feature definitions F1 – F35 plus structural-auxiliary FA – FC.

ID	Name	Definition
F1	Signal power	$\frac{1}{N}\sum_n \lvert r[n]\rvert^2$
F2	Amplitude variance	$\mathrm{Var}(A[n])$
F3	Skewness	$\mathbb{E}[(A-\mu_A)^3]/\sigma_A^3$
F4	Kurtosis	$\mathbb{E}[(A-\mu_A)^4]/\sigma_A^4$
F5	Spectral centroid	$\sum_k k\lvert R[k]\rvert^2/\sum_k \lvert R[k]\rvert^2$
F6	Spectral spread	$\sqrt{\sum_k (k-F_5)^2\lvert R[k]\rvert^2/\sum_k \lvert R[k]\rvert^2}$
F7	IF Std Dev	$\mathrm{std}(\phi'[n])$
F8	IF linear R^2	Linear regression R^2 of $\phi'[n]$ vs. n
F9	IF quadratic R^2	Quadratic regression R^2 of $\phi'[n]$ vs. n
F10	IF curvature	$\mathrm{mean}(\Delta^2\phi'[n])$
F11	IF rate variance	$\mathrm{Var}(\Delta\phi'[n])$
F12	FrFT peak ratio	$\max_\alpha \lvert X_\alpha(u)\rvert^2/E_{\text{total}}$
F13	WVD linearity	R^2 of Wigner-Ville ridge fit in TF plane
F14	Chirp consistency	$1-\mathrm{std}(\phi'[n])/\mathrm{mean}(\lvert\phi'[n]\rvert)$
F15	TF entropy	$-\sum_{t,f} P(t,f)\log P(t,f)$ (STFT spectrogram)
F16	Spectral correlation	$\sum_k R[k]R^*[k+1]/\sum_k \lvert R[k]\rvert^2$
F17	Zero-crossing rate	$(N-1)^{-1}\sum_n \mathbf{1}[r[n]r[n-1]<0]$
F18	PAPR	$\max_n \lvert r[n]\rvert^2/\mathbb{E}[\lvert r[n]\rvert^2]$
F19	Envelope mean	$\mathbb{E}[A[n]]$
F20	Crest factor	$\max_n A[n]/\sqrt{\mathbb{E}[A[n]^2]}$
F21	IF mean	$\mathbb{E}[\phi'[n]]$
F22	Phase variance	$\mathrm{Var}(\phi[n])$
F23	Phase entropy	$-\sum_b p_b\log p_b$ (histogram of $\phi[n]$, 32 bins)
F24	CP correlation	$\lvert\sum_n r[n+N_{\text{CP}}]r^*[n]\rvert/\mathbb{E}[\lvert r\rvert^2]$
F25	Spectral geom. mean	$\exp\big(\frac{1}{N}\sum_k \log\lvert R[k]\rvert\big)$
F26	Spectral flatness	$F_{25}/\mathbb{E}[\lvert R[k]\rvert]$
F27	Spectral occupancy	Fraction of bins: $\lvert R[k]\rvert^2 > 0.1\max\lvert R[k]\rvert^2$
F28	Autocorrelation peak	$\max_{\tau>0}\lvert R_{rr}[\tau]\rvert/R_{rr}[0]$
F29	Spectral entropy	$-\sum_k p_k\log p_k$, $p_k=\lvert R[k]\rvert^2/\sum\lvert R[k]\rvert^2$
F30	Subband variance	$\mathrm{Var}(\bar{P}_b)$, $b=1,\dots,8$ equal-width subbands
F31	Subcarrier regularity	Periodicity score of $\lvert R[k]\rvert^2$ at subcarrier spacing
F32	Periodicity	Peak of normalized autocorrelation of $A[n]$
F33	Cyclic peak	Max of cyclic autocorrelation spectrum
F34	Spectral coherence	$\mathrm{Coh}(R[k],R[k+\Delta])$ averaged over lag Δ
F35	Symbol rate estimation	Dominant period of $\lvert r[n]\rvert^2$ autocorrelation
FA†	Filter sharpness	Ratio of stopband to passband power in subband spectral profile
FB†	Delay correlation	$\lvert\sum_n r[n]r^*[n-D]\rvert/R_{rr}[0]$ at lag $D=N_{\text{CP}}$
FC†	Constellation shape	Normalized variance of $\lvert r[n]\rvert^2$ over symbol grid

Notation: $R[k]$ = DFT of $r[n]$, $A[n]=\lvert r[n]\rvert$, $\phi[n]=\mathrm{unwrap}(\arg(r[n]))$, $\phi'[n]=\Delta\phi[n]$.

†Structural-auxiliary features FA – FC (internal codes `Filter_Sharp`, `Delay_Corr`, `Const_Shape`) rank 10th, 11th, and 13th respectively in Tab. 5. They are extracted from the received block $r[n]$ in fully blind mode and complement F1 – F35. Note: the labels F36, F37, F38 are reserved exclusively for the three novel *channel-aware* features: amplitude fading variance, phase discontinuity, and frequency drift, defined in Sec. 3, Eqs. (13) – (15).

Tab. 3. Multi-algorithm classification performance on the held-out test set (38 features, 70/30 train/test split). Best values are bold.

Classifier	Acc [%]	Prec [%]	Rec [%]	F1 [%]	Train [s]	Inf [ms]	RT?
FC-MLP	**99.09**	**99.10**	**99.09**	**99.09**	970.8	0.021	✓
Ensemble-boost	99.04	99.05	99.04	99.04	157.0	0.023	✓
Random forest[†]	99.02	99.03	99.02	99.03	287.2	0.071[‡]	✓
Decision tree	98.70	98.71	98.70	98.70	3.5	0.001	✓
SVM-linear	97.57	97.62	97.57	97.59	388.0	0.009	✓
SVM-RBF	97.19	97.39	97.19	97.29	2545.7	3.544	×
Ensemble-bag	97.04	97.29	97.04	97.16	47.4	0.024	✓
k-NN	96.94	97.04	96.94	96.99	2.5	1.249	×
Naive Bayes	95.85	96.07	95.85	95.96	3.1	51.681	×

RT? = inference $<$ 1 ms/sample.
[†]RF deployed with 10 features.
[‡]inference time measured on the 10-feature deployment model (0.021 ms at 38 features).

4.5. McNemar Tests

The accuracy differences between all classifier pairs are assessed using the McNemar test with continuity correction [30], which is applied to correct / incorrect vectors per sample. Bonferroni correction for $\binom{9}{2} = 36$ comparisons gives $\alpha_B = 0.05/36 \approx 0.00139$. The full pairwise p-value matrix is shown in Tab. 4. Key findings are as follows:

- FC-MLP, ensemble boost, and random forest are mutually statistically equivalent ($p > \alpha_B$ for all three pairwise comparisons), confirming that these three top classifiers are interchangeable at this operating point,
- All other classifiers are significantly different from the random forest ($p < 0.001$), confirming meaningful accuracy gaps,
- The 10-feature random forest vs. 38-feature random forest gap is not significant ($p > 0.1$), statistically confirming that the compact subset does not lose significant precision.

5. Feature Importance and Selection

Five complementary methods are applied to obtain a robust importance ranking:

1) Random forest permutation importance – OOB accuracy degradation when each feature is permuted,
2) ReliefF – weight-based inter-class vs. intra-class separation (10 nearest neighbors),
3) Chi-square – statistical independence test (10 bins vs. class labels),
4) Mutual information – $I(F;Y) = H(Y) - H(Y|F)$,
5) Fisher score – the ratio of between-class to within-class variance.

The ensemble score shown as Algorithm 2 uses weights $\mathbf{w} = [0.30, 0.175, 0.175, 0.175, 0.175]$, with a higher weight for the importance of random forests due to its direct connection to classifier performance:

$$s_{\text{ensemble}}[i] = \sum_{m=1}^{5} w_m \cdot s_m[i] \,. \tag{16}$$

Algorithm 2 Ensemble feature scoring (leakage-free training fold only)

Require: Feature matrix $\mathbf{F}_{\text{train}} \in \mathbb{R}^{N_{tr} \times 38}$ (training fold only), labels $\mathbf{y}_{\text{train}}$, weights $\mathbf{w}$
Ensure: Ensemble scores $s_{\text{ens}}[i]$, $i = 1, \dots, 38$
for $m = 1$ to 5 **do**
  Compute raw scores $\tilde{s}_m[i]$ with method m on $(\mathbf{F}_{\text{train}}, \mathbf{y}_{\text{train}})$
  Normalize: $\dfrac{s_m[i] = (\tilde{s}_m[i] - \min_i \tilde{s}_m)}{\max_i \tilde{s}_m - \min_i \tilde{s}_m}$
end for
$s_{\text{ens}}[i] = \sum_{m=1}^{5} w_m \cdot s_m[i]$
Return s_{ens}

All scores normalized to $[0, 1]$ before averaging.

5.1. Top Essential Features

Periodicity (rank 1) and subcarrier entropy (rank 2) are the dominant discriminators. OFDM-based waveforms (UFMC, OFDM-IM, GFDM) exhibit a highly regular and periodic subcarrier structure, whereas NOMA (power superposition) and AFDM (chirp modulation) do not. The ensemble score is undefined (NaN), because the chi-square method is not applicable to near-binary indicator features. It is ranked first by all four remaining methods and is assigned rank 1 by consensus.

Channel-aware F36 (`Amp_Fade_Var`) ranks seventh out of 38 (ensemble score 0.360), confirming that explicit fading characterization is highly discriminative. Different waveforms exhibit distinct fading robustness profiles, making channel-induced amplitude variance a powerful and early entering classification cue.

Higher-order statistics (Skewness rank 5, Kurtosis rank 14) distinguish the constant-envelope FBMC-OQAM from the variable-envelope NOMA. Time-frequency entropy (`TF_Entropy`, rank 6) provides robustness to Doppler and multipath through joint time-frequency analysis – a characteristic unavailable to purely spectral or temporal features.

Tab. 4. McNemar test p-values (Bonferroni $\alpha_B = 0.05/36 \approx 0.00139$).

	k-NN	SVM-linear	SVM-RBF	Decision tree	Random forest	Naive Bayes	Ensemble-boost	Ensemble-bag	FC-MLP
k-NN	–	**0.000**	0.001†	**0.000**	**0.000**	**0.000**	**0.000**	0.218	**0.000**
SVM-linear		–	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
SVM-RBF			–	**0.000**	**0.000**	**0.000**	**0.000**	0.043	**0.000**
Decision tree				–	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**
Random forest					–	**0.000**	0.649	**0.000**	0.039
Naive Bayes						–	**0.000**	**0.000**	**0.000**
Ensemble-boost							–	**0.000**	0.082
Ensemble-bag								–	**0.000**
FC-MLP									–

Bold entries are statistically significant ($p < \alpha_B$). FC-MLP, ensemble-boost, and random forest are mutually non-significant (blue).
†Exact $p = 0.00148 > \alpha_B \approx 0.00139$: not significant after Bonferroni correction.
All values verified against `mcnemar_tests.mat`.

Tab. 5. Top 15 features ranked by weighted ensemble score (RF 30% + ReliefF 17.5% + χ^2 17.5% + MI 17.5% + Fisher 17.5%).

Rank	Feature	Score	Type
1	Period	–†	Spectral
2	Subcar. Entropy	0.825	Temporal
3	Var. Amp	0.515	Spectral
4	Spec. Occup	0.502	Spectral
5	Skewness	0.428	Statistical
6	TF Entropy	0.405	Statistical
7	**Amp. Fade Var (F36)**	**0.360**	**Channel-aware**
8	Spec. Coher	0.353	Spectral
9	CP Corr	0.326	Waveform-specific
10	Filter Sharp	0.321	Spectral
11	Delay Corr	0.315	Statistical
12	Subcar. Reg	0.306	Spectral
13	Const. Shape	0.275	Statistical
14	Kurtosis	0.275	Statistical
15	IF Std	0.272	Inst.-freq.

† Score undefined: χ^2 returns NaN for period (near-binary indicator incompatible with frequency-bin discretization). The period is ranked first by all the other four methods; it is assigned rank 1 by consensus. Yellow rows: optimal 10-feature deployment subset. F36 (`Amp_Fade_Var`) is the novel channel-sensitive feature. Scores verified against `feature_importance.mat`.

5.2. Channel-aware Feature Ablation Study

Table 6 shows the impact of removing channel-aware features. F36 is most critical for fading (–1.4% when removed alone), F37 for Doppler and CFO conditions (–1.1% alone under high Doppler). F38 (rank 36/38, score 0.103) contributes primarily under extreme pre-synchronization CFO scenarios and does not enter the 10-feature subset.

Tab. 6. Ablation study: accuracy in percent of the 10-feature random forest with/without channel-aware features. Verified against `per_channel_confusion.mat`.

Channel	All	F36	F37	F38	None
AWGN	99.8	99.8	99.8	99.8	99.8
Rayleigh/Rician	98.5	97.1	98.3	98.4	96.4
Frequency offset	97.9	97.8	96.8	97.7	96.6
Low Doppler	98.2	98.0	97.7	97.9	96.9
High Doppler	96.8	96.6	95.7	96.3	94.2
Freq.-selective	97.4	97.1	97.3	97.3	96.8
Timing offset	98.1	98.0	97.4	98.0	97.2
Double-selective	96.1	95.5	95.3	95.6	93.0
Mean	97.9	97.5	97.3	97.6	96.4

6. Feature Reduction Analysis

The random forest achieves a performance rate of 98.90% with 10 features and 99.05% with 15, a gap of 0.15 pp. The 10-feature configuration is within 0.12 percentage points (pp) of the full-feature baseline at 74% dimensionality reduction. Strong performance is achieved even with five features (95.38 – 97.76%), well above the 12.5% chance level for 8 classes, and all classifiers already exceed 95%.

There is a plateau beyond 10 features for the random forest, where the accuracy fluctuates by no more than 0.13 pp from 10 to 38 features. Ensemble-boost shows a 0.22 pp gap at 10 features (98.82% vs. 99.04% at 38 features). The plateau of < 0.15 pp is reached from 15 features onward for all ensemble methods.

All four classifiers converge to near-optimal at 10 features, validating classifier-independent feature selection. FC-MLP tracks random forest and ensemble boost closely across all feature counts.

Tab. 7. Classification accuracy (in percent) vs. feature count for four classifiers.

Feature count	Ens.-boost	Rand. forest	Dec. tree	FC-MLP
5	97.58	97.76	95.38	97.13
10	98.82	98.90	97.52	98.82
15	98.97	99.05	98.10	99.02
20	99.02	99.06	98.43	99.05
25	99.03	99.05	98.78	99.08
30	99.03	99.06	98.70	99.04
38	99.04	99.02	98.70	99.09

RF: <0.17 pp gain beyond 10 features (38F baseline 99.02%). Ens.-Boost: 0.22 pp gap at 10F; plateau <0.15 pp from 15F onward. All classifiers use the same subset of ranked features and identical 70/30 train/test partition. Optimal row (10 features) shaded. Decision tree accuracy recovers more slowly due to overpruning. Values verified against `feature_reduction.mat`.

The 10-feature (10F) configuration yields: (i) 74% reduction in feature extraction operations; the costliest feature, WVD (F13, $\mathcal{O}(N^2 \log N)$, rank 15), is eliminated, reducing total extraction time from 18.7 ms to 2.1 ms per block at the 8 kHz simulation rate – see Sec. 8.3 for wideband projections; (ii) 73.7% memory reduction from 152 to 40 bytes per feature vector; (iii) 0.071 ms inference time per segment on a standard Intel Core i7-8550U, 2.4 GHz processor, enabling $>$ 14 000 segments/s.

7. Cross-validation and Robustness Analysis

7.1. K-Fold Cross-validation

The 10-feature random forest model undergoes 5-fold stratified cross-validation: mean accuracy 98.54%, $\sigma = 0.05\%$, Wilson 95% CI [98.49%, 98.59%] [31] and coefficient of variation 0.05%. The Wilson interval is used in preference to the normal approximation, which is inaccurate near $p = 1$. The narrow interval confirms the exceptional generalization stability. Feature selection is performed independently within each fold's training partition, preserving strict leakage-free evaluation throughout.

7.2. Per-Channel Performance

Using the 10-feature ensemble-boost model on the nine custom channels (all values verified against `per_channel_confusion.mat`):

- AWGN: 99.96%,
- Rayleigh fading: 98.87% (F36 maintains discrimination under amplitude fluctuations),
- Rician fading: 99.98%,
- Frequency offset: 99.42% (cyclostationary features robust to constant CFO; residual confusion in AFDM/OTFS pair),
- Timing offset: 99.92%,
- Low Doppler (30 km/h): 98.93%,
- High Doppler (120 km/h): 99.04%,
- Frequency-selective: 99.98% (multipath creates distinctive temporal correlation aiding UFMC identification),
- Double selective time-frequency: 99.02% without F36 – F38, this drops to 93.0%, as reported in the ablation study in Tab. 6.

3GPP TDL standard channels (verified against `per_channel_confusion.mat`):

- TDL-A (30 ns delay spread, indoor/dense urban): 99.96% confirming transferability of selected features to standard profiles,
- TDL-B (100 ns, urban macro): 99.95%,
- TDL-C (300 ns, rural macro): 99.97% – the largest standard profile delay spread, yet performance remains consistently high, confirming robustness of the selected feature set.

Tab. 8. Confusion matrix under the Rayleigh fading channel.

True/Pred	UFMC	GFDM	FBMC	NOMA	OFDM-IM	OTFS	ODDM	AFDM
UFMC	**2100**	0	0	0	0	0	0	0
GFDM	0	**2100**	0	0	0	0	0	0
FBMC	0	0	**2007**	93	0	0	0	0
NOMA	0	0	56	**2044**	0	0	0	0
OFDM-IM	0	0	0	0	**2082**	0	18	0
OTFS	0	0	0	0	0	**2100**	0	0
ODDM	0	0	0	0	23	0	**2077**	0
AFDM	0	0	0	0	0	0	0	**2100**

10-feature ensemble-boost model; 2100 test samples per class. The bold diagonal entries are per class recall.

Table 8 shows the 8 × 8 confusion matrix generated by the 10-feature model under the Rayleigh fading channel. UFMC, GFDM, OTFS and AFDM achieve perfect classification (2100 / 2100). The primary confusion occurs between FBMC and NOMA (93 FBMC samples misclassified as NOMA, 56 NOMA samples misclassified as FBMC), which is attributable to their overlapping envelope statistics under amplitude-varying fading. A secondary confusion cluster appears between OFDM-IM and ODDM (18 and 23 samples, respectively), reflecting their shared index-modulation and spreading characteristics.

7.3. Parametric Robustness Analysis

SNR sensitivity (10-feature random forest model): at 5 dB is 84.5%; at 0 dB is 93.2%; at +5 dB is 97.6%. In $\geqslant$ 10 dB range it equals $\geqslant$ 98.4%. Below –5 dB, signal averaging or diversity combining is recommended.

Doppler sensitivity: 0 km/h: 99.4%; 30 km/h: 98.2%; 60 km/h: 97.5%; 120 km/h: 96.8%. Performance degrades gracefully across the full range from urban to highway. CFO sensitivity: 150 Hz (post-synchronization, $\approx$ 0.063 ppm):

Tab. 9. Comparison with state-of-the-art AMC methods. This work adds a FC-MLP deep learning baseline and 3GPP TDL standard channels.

Reference	Waveforms	Channels	Acc [%]	Method
[32]	11 conv.	AWGN only	94.20	HO cumulants
[10]	5 NG	Simple	95.00	Deep ResNet+PCA
[33]	Conv.	AWGN+ Rician	96.80	CNN-LSTM
[34]	Conv.	AWGN	95.60	Deep CNN
This work (FC-MLP, 38F)	8 NG	12 channels	99.09	MLP 256-128-64
This work (RF, 10F)	8 NG	12 channels	98.90	RF, 10 features

74% reduction in features, 0.071ms / segment, capable of real-time.
NG = next-generation waveform.

97.9%; 1 kHz: 96.9%; 2.4 kHz (≈ 1 ppm at 2.4 GHz, pre-synchronization): 96.3%; 5 kHz: 93.1%. The 2.4 kHz case remains at 96.3%, confirming practical robustness for consumer-grade oscillator mismatch. Timing offset sensitivity: 0 samples: 99.8%; 5 samples: 98.1%; 10 samples: 97.9%; 15 samples: 96.6% (faster degradation as CP is partially corrupted).

8. Discussion

8.1. Comparison with Other Work

The proposed framework achieves a 2.09 – 4.89 percentage point improvement over the state-of-the-art solution on comparable or simpler waveform sets – Tab. 9. The authors of [32] used higher-order cumulants for conventional PSK/QAM and their approach does not extend to next-generation multicarrier waveforms. In [10], a deep residual network achieving approximately 95% accuracy is proposed in five 5G waveforms, but the evaluation excluded delay Doppler waveforms (OTFS, ODDM, AFDM) and compound channel conditions. Frameworks described in [33], [34] operate on conventional modulation schemes under simpler channel models.

Critically, the deep learning baseline trained on the same 38 characteristics achieves 99.09%, which is the highest single model precision score and is statistically equivalent to that of ensemble boost and random forest (McNemar $p > \alpha_B$ for all three pairwise comparisons). This result demonstrates two complementary points: (i) the hand-crafted features are rich enough to support top-tier deep learning classification, and (ii) structured ensemble classifiers that explicitly exploit feature interactions via tree splits achieve performance equivalent to that of generic MLP architectures on this problem, while offering physical interpretability.

The dominant discriminators, i.e., periodicity, subcarrier entropy, F36, are explicitly identified – a feat that end-to-end black-box approaches cannot offer.

8.2. Practical CR Deployment Considerations

The 10-feature compact model is specifically designed for the resource envelope of real CR hardware. It requires $<$ 40 KB total memory (40-byte feature vectors plus 35 KB for random forest model), placing it within the capabilities of ARM Cortex-M7 microcontrollers, FPGAs, software-defined radio platforms (USRP, HackRF, LimeSDR), UAV-mounted spectrum sensors and battery-powered secondary IoT nodes. With 0.071 ms inference time per segment, the identifier imposes negligible latency overhead on the CR sensing decision cycle and may comfortably run in parallel with other CR functions such as channel estimation and power control.

For access to the dynamic CR spectrum, the identification latency directly determines how quickly a secondary user can react to the transmission of a primary user. At 0.071 ms per block, the system supports $>$ 14 000 spectrum sensing decisions per second within the millisecond reaction times required by 5G NR and anticipated by 6G MAC protocols. The 40-byte feature vector also lends itself to cooperative CR sensing – secondary nodes can broadcast their feature vectors to a fusion center for ensemble decision-making without significant control channel overhead.

8.3. Scalability to Wideband CR Systems

All frequency domain features are normalized by the Nyquist frequency, and all temporal features use symbol-period units rather than absolute sample counts. Therefore, the features are scale-invariant by design. A CR receiver observing a 5G NR signal at 100 MHz bandwidth would extract identical normalized subcarrier regularity and spectral occupancy ratios as in the 8 kHz experiments reported here.

It is important to be explicit about the current scope. The experiments use f_s = 8 kHz as a reduced-scale proof-of-concept that preserves all waveform structural properties (subcarrier relationships, CP ratios, Doppler-to-bandwidth ratios) while remaining computationally tractable for a comprehensive 201 600 signal, 12 channel, 9 classifier study. The extrapolation to wideband CR bandwidths is analytically motivated, but is not yet hardware validated.

Feature extraction time scales as follows: at 8 kHz, the dominant cost is WVD ($\mathcal{O}(N^2 \log N)$), giving 18.7 ms per block for the full 38-feature set and 2.1 ms for the 10-feature subset (WVD eliminated). At 100 MHz, representative of a wideband CR receiver, symbol-period windows are proportionally shorter ($\sim$150$\times$ fewer samples for the same number of symbols), so extraction time is expected to reduce to $\approx$ 30 µs per block. Hardware validation at 5G NR bandwidths (e.g., on a USRP B210 or ZCU111 RFSoC operated as a CR front end) remains a priority for future work and is listed as a limitation below.

8.4. Limitations and Future Directions

All experiments use f_s = 8 kHz. Scale invariance is analytically guaranteed by feature normalization but not yet confirmed by hardware measurement at 5G NR bandwidths on a real CR platform. USRP or RFSoC validation is the most urgent open item.

Waveform block lengths range from 72 (OTFS) to 512 samples (GFDM, FBMC). In a CR receiver with no a priori knowledge of the waveform, block segmentation must be fixed, and longer blocks yield more stable feature estimates. Per-waveform confusion analysis (Tab. 8) reveals that FBMC and NOMA sharing overlapping envelope statistics under fading are the primary confusion pair. Future work should investigate adaptive block-length strategies or multiblock averaging.

The current framework assumes that one of the eight known waveforms is always present. Confidence-based rejection (maximum posterior class probability < 0.7) can flag unknown or novel waveforms; preliminary experiments correctly rejected 78% of artificially injected out-of-distribution signals. Open-set CR sensing remains an open problem.

A CR receiver's own front-end introduces phase noise of –80 dBc/Hz and I/Q imbalance (0.1 – 0.5 dB amplitude mismatch), causing an additional accuracy loss of up to 1.4% and 0.8%, respectively, in supplementary Saleh PA model experiments. The 16-QAM substitution causes only a degradation of 0.3 to 0.8%, confirming that waveform-structural features are largely modulation order independent, an important property for blind CR sensing.

Sinc-interpolation-based offsets of 0.1 – 0.5 samples cause an additional degradation of 0.4 – 1.2%, mainly affecting CP correlation F24. A CR receiver with multiple antennas could exploit MIMO spatial signatures as additional discriminative features.

Future directions are as follows:

- Hardware CR validation: Deploy the 10-feature identifier on a USRP B210 or ZCU111 RFSoC operating as a CR secondary node, and confirm real-time throughput and scale invariance at 5G NR bandwidths.
- Raw I/Q deep learning for CR: 1D-ResNet or transformer trained on raw I/Q, with the 38-feature model as an interpretable reference for attention visualization and CR explainability requirements.
- Transfer and meta-learning: A few-shot adaptation of the CR identifier to novel 6G waveform candidates, using the 10-feature space as a compact domain-shift-robust representation.
- Hierarchical CR identification: Two-stage approach (waveform family, then specific waveform) to resolve FBMC/NOMA and OFDM-IM/ODDM confusion, reducing misclassification-driven interference in CR systems.
- Joint identification and parameter estimation: Simultaneous waveform identification and symbol rate/modulation order estimation to fully characterize the primary user signal for CR resource allocation.
- Cooperative CR sensing: Distributed identification where multiple CR secondary nodes share compact feature vectors for fusion center ensemble decisions, exploiting spatial diversity.
- Open-set CR sensing: The detection of novelty of a single class SVM or isolation forest in the 10-feature space to flag unknown or emerging waveforms not in the training set.

9. Conclusions

This paper presented a framework for the identification of cognitive radio waveforms for eight 5G-Advanced/6G waveform candidates under realistic compound channel impairments, addressing the fundamental CR requirement of blind, non-cooperative waveform sensing. The study was extended with three 3GPP TDL standard profiles, an FC-MLP deep learning baseline, leakage-free feature selection, and McNemar statistical validation. The following findings have been verified based on 201 600 signals under twelve channel conditions and with seven SNR levels:

1) FC-MLP achieves 99.09% waveform identification accuracy with 38 features 2.09 – 4.89 pp above prior state-of-the-art solution in comparable or simpler waveform sets. FC-MLP, ensemble boost (99.04%) and random forest (99.02%) are mutually statistically equivalent (McNemar $p > \alpha_B$ for all pairwise comparisons), giving CR system designers flexibility in choosing between interpretable ensemble methods and deep learning.
2) The hand-crafted 38-feature set fully supports deep-learning-based CR identification. FC-MLP reaches 99.09% on the same features, while ensemble tree methods achieve statistically equivalent performance with greater interpretability, a key property for CR regulatory and cognitive transparency requirements.
3) Compact CR deployment enabled by 10-feature random forest at 98.90% within 0.12% of the full-feature baseline (McNemar $p > 0.1$), with a 74% reduction in feature extraction cost and 0.071 ms per segment inference, satisfying the real-time latency budget for dynamic CR spectrum access.
4) Novel channel-aware features are critical for the robustness of CR. Ablation studies confirm an accuracy improvement of up to 3.1% under double-selective fading (93.0% → 96.1%), i.e. the most adversarial CR propagation scenario, with F36 (amplitude fading variance) ranking 7th among 38 features and entering the optimal 10-feature CR deployment subset. F38 (frequency drift, rank 36/38) contributes primarily under extreme CFO conditions and does not enter the 10-feature subset.
5) Robustness throughout the entire CR channel envelope. Identification accuracy boosted from 98.87% (Rayleigh fading) to 99.98% (AWGN and Rician fading) on custom channels and from 99.95% (TDL-B) to 99.97% (TDL-C) on 3GPP standard profiles, confirming direct transferability to 5G NR CR deployments. Per-waveform confusion analysis identifies FBMC/NOMA

and OFDM-IM/ODDM as the primary confusion pairs under Rayleigh fading, providing actionable guidance for CR system design.

6) Cross-validation confirms generalization stability: 98.54% mean, Wilson 95% CI [98.49%, 98.59%], σ = 0.05% in 5 folds. All feature selection is performed within training folds, ensuring genuine out-of-sample performance estimates suitable for CR deployment decisions.
7) Algorithm independent compact sensing. The top 10 leakage-free feature subsets are transferred across random forest, ensemble-boost, decision tree, and FC-MLP, enabling CR platform designers to select their preferred classifier without redesigning the sensing front-end.

The proposed framework enables practical cognitive radio waveform identification (40-byte feature vectors, 0.071 ms latency, $>$ 14 000 sensing decisions per second) while providing a physical insight into discriminative mechanisms: periodicity, subcarrier entropy, and channel-aware F36 that CR systems can leverage for adaptive spectrum management. Hardware validation of the compact 10-feature identifier on a USRP or RFSoC CR platform at 5G NR bandwidths is the most critical open item for the transition from simulation to real CR deployment.

References

[1] O.A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of Automatic Modulation Classification Techniques: Classical Approaches and New Trends", *IET Communications*, vol. 1, pp. 137–156, 2007 (https://doi.org/10.1049/iet-com:20050176).

[2] P.C. Sahu *et al.*, "Cloud-enabled Automatic Modulation Classification Using Deep Feature Fusion and Moth-flame Optimized Elm Approach", *Scientific Reports*, vol. 16, art. no. 1061, 2026 (https://doi.org/10.1038/s41598-025-30753-4).

[3] S. Chen and J. Zhao, "The Requirements, Challenges, and Technologies for 5G of Terrestrial Mobile Telecommunication", *IEEE Communications Magazine*, vol. 52, pp. 36–43, 2014 (https://doi.org/10.1109/MCOM.2014.6815891).

[4] P. Li, J. Fan, and J. Wu, "Exploring the Key Technologies and Applications of 6G Wireless Communication Network", *iScience*, vol. 28, art. no. 112281, 2025 (https://doi.org/10.1016/j.isci.2025.112281).

[5] W. Saad, M. Bennis, and M. Chen, "A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems", *IEEE Network*, vol. 34, pp. 134–142, 2020 (https://doi.org/10.1109/MNET.001.1900287).

[6] W. Jiang, B. Han, M.A. Habibi, and H.D. Schotten, "The Road Towards 6G: A Comprehensive Survey", *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, 2021 (https://doi.org/10.1109/OJCOMS.2021.3057679).

[7] L. Zhang *et al.*, "Subband Filtered Multi-carrier Systems for Multiservice Wireless Communications", *IEEE Transactions on Wireless Communications*, vol. 16, pp. 1893–1907, 2017 (https://doi.org/10.1109/TWC.2017.2656904).

[8] R. Hadani *et al.*, "Orthogonal Time Frequency Space Modulation", *IEEE Wireless Communications and Networking Conference (WCNC)*, San Francisco, USA, 2017 (https://doi.org/10.1109/WCNC.2017.7925924).

[9] M. Wen, X. Cheng, and L. Yang, *Index Modulation for 5G Wireless Communications*, Springer, 164 p., 2017 (https://doi.org/10.1007/978-3-319-51355-3).

[10] H.B. Chikha, A. Alaerjan, and R. Jabeur, "Automatic Classification of 5G Waveform-modulated Signals Using Deep Residual Networks", *Sensors*, vol. 25, art. no. 4682, 2025 (https://doi.org/10.3390/s25154682).

[11] 3GPP, "Study on Channel Model for Frequencies from 0.5 to 100 GHz", Technical Report TR 38.901, 3rd Generation Partnership Project, 2022.

[12] W.C. Jakes (ed.), *Microwave Mobile Communications*, Wiley-IEEE Press, 656 p., 1994 (ISBN: 9780780310698).

[13] N. Michailow *et al.*, "Generalized Frequency Division Multiplexing for 5th Generation Cellular Networks", *IEEE Transactions on Communications*, vol. 62, pp. 3045–3061, 2014 (https://doi.org/10.1109/TCOMM.2014.2345566).

[14] M. Matthe, N. Michailow, I. Gaspar, and G. Fettweis, "Influence of Pulse Shaping on Bit Error Rate Performance and out of Band Radiation of Generalized Frequency Division Multiplexing", *2014 IEEE International Conference on Communications Workshops (ICC)*, Sydney, Australia, 2014 (https://doi.org/10.1109/ICCW.2014.6881170).

[15] M. Bellanger *et al.*, "FBMC Physical Layer: A Primer", PHYDYAS, 2010.

[16] M. Besseghier, A.B. Djebbar, and E. Kofidis, "Joint CFO and Highly Frequency Selective Channel Estimation in FBMC/OQAM Systems", *Digital Signal Processing*, vol. 128, art. no. 103629, 2022 (https://doi.org/10.1016/j.dsp.2022.103629).

[17] Y. Liu *et al.*, "Nonorthogonal Multiple Access for 5G and Beyond", *Proceedings of the IEEE*, vol. 105, pp. 2347–2381, 2017 (https://doi.org/10.1109/JPROC.2017.2768666).

[18] Z. Ding *et al.*, "Application of Non-orthogonal Multiple Access in LTE and 5G Networks", *IEEE Communications Magazine*, vol. 55, pp. 185–191, 2017 (https://doi.org/10.1109/MCOM.2017.1500657CM).

[19] M. Besseghier, A. Zouggaret, S. Ghouali, and A.B. Djebbar, "Performance Degradation Analysis of Downlink NOMA Systems Under Carrier Frequency Offsets", *IEEE Transactions on Consumer Electronics*, vol. 71, pp. 2508–2516, 2025 (https://doi.org/10.1109/TCE.2025.3571676).

[20] E. Basar, U. Aygolu, E. Panayirci, and H.V. Poor, "Orthogonal Frequency Division Multiplexing with Index Modulation", *IEEE Transactions on Signal Processing*, vol. 61, pp. 5536–5549, 2013 (https://doi.org/10.1109/TSP.2013.2279771).

[21] E. Basar *et al.*, "Index Modulation Techniques for Next-generation Wireless Networks", *IEEE Access*, vol. 5, pp. 16693–16746, 2017 (https://doi.org/10.1109/ACCESS.2017.2737528).

[22] M. Besseghier, S. Ghouali, and A.B. Djebbar, "Optimized Greedy Detection for OFDM-IM Systems", *IEEE Communications Letters*, vol. 27, pp. 2034–2037, 2023 (https://doi.org/10.1109/LCOMM.2023.3283032).

[23] P. Raviteja, K.T. Phan, Y. Hong, and E. Viterbo, "Interference Cancellation and Iterative Detection for Orthogonal Time Frequency Space Modulation", *IEEE Transactions on Wireless Communications*, vol. 17, pp. 6501–6515, 2018 (https://doi.org/10.1109/TWC.2018.2860011).

[24] G.D. Surabhi, R.M. Augustine, and A. Chockalingam, "On the Diversity of Uncoded OTFS Modulation in Doubly-dispersive Channels", *IEEE Transactions on Wireless Communications*, vol. 18, pp. 3049–3063, 2019 (https://doi.org/10.1109/TWC.2019.2909205).

[25] G.D. Surabhi and A. Chockalingam, "Low-complexity Linear Equalization for OTFS Modulation", *IEEE Communications Letters*, vol. 24, pp. 330–334, 2020 (https://doi.org/10.1109/LCOMM.2019.2956709).

[26] Z. Wei *et al.*, "Orthogonal Time-frequency Space Modulation: A Promising Next-generation Waveform", *IEEE Wireless Communications*, vol. 28, pp. 136–144, 2021 (https://doi.org/10.1109/MWC.001.2000408).

[27] A. Bemani, N. Ksairi, and M. Kountouris, "Affine Frequency Division Multiplexing for Next Generation Wireless Communications", *IEEE Transactions on Wireless Communications*, vol. 22, pp. 8214–8229, 2023 (https://doi.org/10.1109/TWC.2023.3260906).

[28] A. Bouguila, A. Khlifi, and M. Yahia, "Performance Analysis of OFDM, OTFS, and AFDM Modulation Over Doubly Selec-

tive Channels", *2025 IEEE International Multi-Conference on Smart Systems & Green Process (IMC-SSGP)*, Hammamet, Tunisia, 2025, (https://doi.org/10.1109/IMC-SSGP67001.2025.11473997).

[29] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 800 p., 2016 (https://doi.org/10.1007/s10710-017-9314-z).

[30] Q. McNemar, "Note on the Sampling Error of the Difference Between Correlated Proportions or Percentages", *Psychometrika*, vol. 12, pp. 153–157, 1947 (https://doi.org/10.1007/BF02295996).

[31] E.B. Wilson, "Probable Inference, the Law of Succession, and Statistical Inference", *Journal of the American Statistical Association*, vol. 22, pp. 209–212, 1927 (https://doi.org/10.2307/2276774).

[32] A. Swami and B.M. Sadler, "Hierarchical Digital Modulation Classification Using Cumulants", *IEEE Transactions on Communications*, vol. 48, pp. 416–429, 2000 (https://doi.org/10.1109/26.837045).

[33] Z. Zhang *et al.*, "Automatic Modulation Classification Using CNN-LSTM Based Dual-stream Structure", *IEEE Transactions on Vehicular Technology*, vol. 69, pp. 13521–13531, 2020 (https://doi.org/10.1109/TVT.2020.3030018).

[34] F. Meng, P. Chen, L. Wu, and X. Wang, "Automatic Modulation Classification: A Deep Learning Enabled Approach", *IEEE Transactions on Vehicular Technology*, vol. 67, pp. 10760–10772, 2018 (https://doi.org/10.1109/TVT.2018.2868698).

Abdelkader Horch, Ph.D.
https://orcid.org/0009-0001-4382-4996
E-mail: a.horch@univ-mascara.dz
University of Mustapha Stambouli, Mascara, Algeria
https://www.univ-mascara.dz

Mokhtar Besseghier, Ph.D.
https://orcid.org/0000-0003-3522-1332
E-mail: m.besseghier@univ-mascara.dz
University of Mustapha Stambouli, Mascara, Algeria
https://www.univ-mascara.dz

Samir Ghouali, Ph.D.
https://orcid.org/0000-0002-3653-873X
E-mail: s.ghouali@univ-mascara.dz
University of Mustapha Stambouli, Mascara, Algeria
https://www.univ-mascara.dz

Abderrahmane Louni, Ph.D.
https://orcid.org/0009-0007-6401-4363
E-mail: a.louni@univ-mascara.dz
University of Mustapha Stambouli, Mascara, Algeria
https://www.univ-mascara.dz