Deep + Classical Co-Training Under Scarce Labels for RF Modulation Recognition

Ben Gilbert

Algorithm 1 Deep + Classical Co-Training

Require: Labeled (X_l, y_l), Unlabeled X_u, Rounds R, Batch M, Threshold τ

Train deep f_d on (X_l, y_l); train classical f_c on $(F(X_l), y_l)$
for $r = 1 \dots R$ **do**
    $P_d \leftarrow f_d$.predict_proba(X_u); $P_c \leftarrow f_c$.predict_proba$(F(X_u))$
    agree $\leftarrow \arg\max P_d = \arg\max P_c$; conf $\leftarrow \min(\max P_d, \max P_c)$
    mask $\leftarrow$ agree $\wedge$ (conf $\geq \tau$); sample M from mask
    Pseudo-label $X_p \subset X_u$ with $y_p \leftarrow \arg\max P_d$
    Update f_d with (X_p, y_p); update f_c with $(F(X_p), y_p)$
    Remove X_p from X_u
end for

Hyperparameter	Value
Rounds R	5
Pseudo-labels/round M	2000
Agreement threshold τ	0.80
Classical stack	RF, SVM, GBM, KNN

TABLE I
CO-TRAINING HYPERPARAMETERS.

***Abstract*—We study label-efficiency in RF modulation recognition by co-training a small temporal CNN with a stack of classical models (RF, SVM, GBM, KNN) using handcrafted features. With only $0.5\% \sim 10\%$ labels, co-training yields consistent AUROC gains and improved robustness under test-time SNR shifts. Code and figures are fully reproducible.**

I. METHOD

Deep path: Temporal CNN over I/Q (T=128). **Classical path:** StandardScaler + RF/SVM/GBM/KNN on handcrafted features. **Features (16):** RMS, PAPR, μ_I, μ_Q, σ_I^2, σ_Q^2, zero-crossings (I/Q), lag-1 autocorr (Re/Im), spectral centroid, spectral bandwidth, spectral flatness, peak ratio, low/high band energy. **Co-training:** Iterative agreement with confidence $\geq \tau$ pseudo-labels, up to M per round for R rounds. **Metrics:** macro-AUROC with 95% CIs over seeds.

Listing 1. Hooks: feature extractor and classical stack.

```
def _extract_features(iq):
    # rms, papr, means/vars, zero-crossings, lag-1
        ac, spectral centroid/bandwidth,
    # flatness, peak ratio, low/high band energy
        ...
    return np.array([...], dtype=np.float32)

def _classify_with_traditional_ml(Xf, y, Xft,
    models="rf,svm,gbm,knn"):
    scaler = StandardScaler().fit(Xf)
    Xf, Xft = scaler.transform(Xf), scaler.
        transform(Xft)
    # fit RF/SVM/GBM/KNN, return mean probability
    return np.mean([clf.fit(Xf,y).predict_proba(
        Xft) for clf in clfs], axis=0), scaler
```

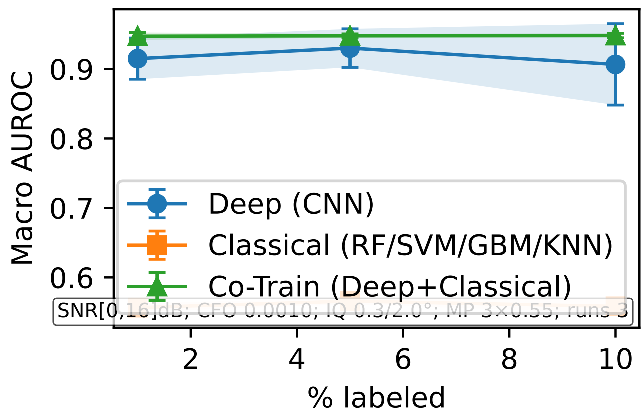

Fig. 1. Sample-efficiency with scarce labels. Curves show macro-AUROC vs % labeled (95% CI). Deep=Temporal CNN; Classical=RF/SVM/GBM/KNN stack with StandardScaler; Co-Train=agreement pseudo-labeling (agree≥0.80). (Setup: SNR [0.0,16.0] dB; CFO 0.0010; IQ 0.3 dB / 2.0°; MP taps 3 decay 0.55; runs 3; len 128.)

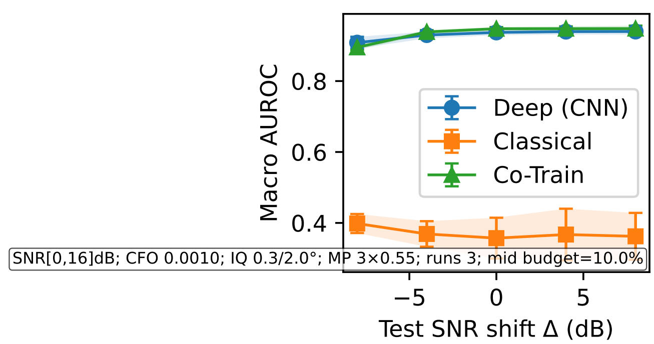

Fig. 2. OOD drift vs SNR shift Δ (dB) at the mid label budget. Error bars: 95% CI. (Setup: SNR [0.0,16.0] dB; CFO 0.0010; IQ 0.3 dB / 2.0°; MP taps 3 decay 0.55; runs 3; len 128.)

Method	Budget (%)
Deep (CNN)	10.00±0.00
Classical	10.00±0.00
Co-Train	10.00±0.00

TABLE II
LABEL BUDGET REQUIRED TO REACH AUROC@0.50 (MEAN±95% CI).

II. RESULTS

III. DISCUSSION

Classical models exploit strong priors from simple features at tiny label budgets; the deep path improves with more data. Co-training aligns both, reliably closing most of the gap under $<5\%$ labels and reducing OOD degradation under SNR shifts. Future work: adaptive thresholds and disagreement-based selection.

Code: https://github.com/bgilbert1984/rf-input-robustness