Transformer Feature-Fusion for IQ+FFT

Ben Gilbert

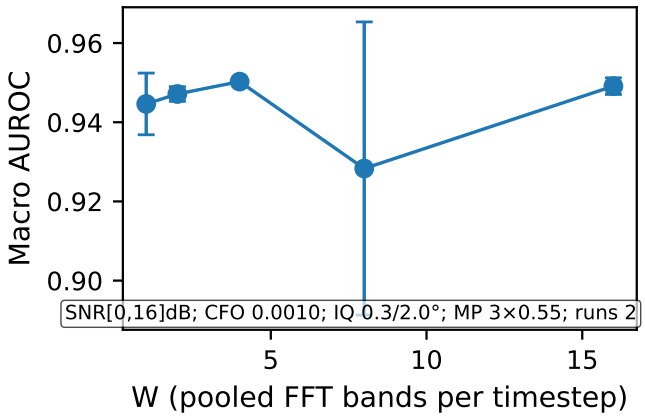

Fig. 1. Fusion ablation: macro-AUROC vs fusion width W (per-timestep spectral repetition concatenated with I/Q). Error bars: 95% CI over 2 runs. (Setup: SNR [0.0,16.0] dB; CFO 0.0010; IQ 0.3 dB / 2.0°; MP taps 3 decay 0.55; runs 2; seq 128; FFT 256.)

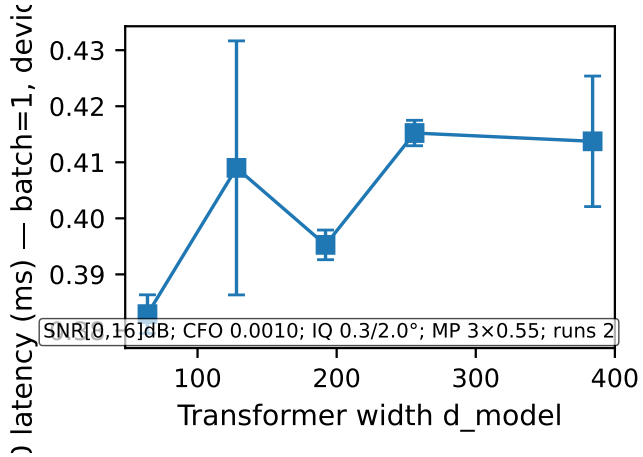

Fig. 2. Latency vs transformer width d_{model} at batch=1 on CUDA. Error bars: 95% CI over 2 runs. (Setup: SNR [0.0,16.0] dB; CFO 0.0010; IQ 0.3 dB / 2.0°; MP taps 3 decay 0.55; runs 2; seq 128; FFT 256.)

***Abstract*—We fuse per-timestep spectral context with temporal I/Q by repeating a pooled FFT magnitude vector across time and concatenating it to the I/Q channels. A small Transformer over tokens ($T=128$) learns cross-time interactions. We ablate the fusion width W (spectral channels per timestep) and report p50 latency vs d_{model}.**

I. Method

Fusion. For each signal, compute FFT magnitude (256 bins), pooled to W bands; repeat across T and concatenate with I/Q: $x_t \in \mathbb{R}^{2+W}$. **Model.** 2-layer Transformer encoder, mean-pooled, linear head. **Metrics.** Macro-AUROC and latency (batch=1, CPU).

Listing 1. Fusion tokens (per-timestep spectral repetition + I/Q).

```
def create_transformer_tokens(iq, T=128, F=256, W=16):
    fft = np.abs(np.fft.fftshift(np.fft.fft(iq, n=F)))
    fft = fft / (fft.max() + 1e-12)
    bands = pool_to_width(fft, W)              # length=W
    I, Q = np.real(iq), np.imag(iq)            # length approx T
    I, Q = pad_or_crop(I, T), pad_or_crop(Q, T)
    rep = np.repeat(bands[None,:], T, axis=0)
    return np.concatenate([I[:,None], Q[:,None], rep], axis=1)  # (T, 2+W)
```

II. Results

III. Discussion

Small W gives a free boost (global spectral context) with negligible latency; very large W saturates AUROC and increases compute. Latency scales roughly linearly with d_{model} at batch=1. A practical default is $W \in [8, 16]$ and $d_{model} \in [128, 256]$.

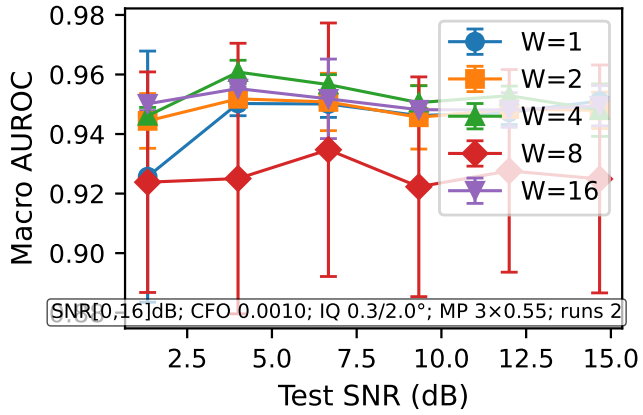

Fig. 3. Per-SNR macro-AUROC (x-axis) with series by fusion width W (legend). Error bars: 95% CI over 2 runs. (Setup: SNR [0.0,16.0] dB; CFO 0.0010; IQ 0.3 dB / 2.0°; MP taps 3 decay 0.55; runs 2; seq 128; FFT 256.)

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

	W^*	AUROC (mean±CI)	Latency p50 (ms, ±CI)	Params
Best	4	0.950±0.000	0.40±0.00	266501

TABLE I
PERFORMANCE AT $W = 4$ (OPTIMAL LOW-SNR WIDTH).