

Mamba による WiFi CSI 人物認証：NTU-Fi HumanID における比較評価

荻山 湊（東京大学 工学部システム創成学科 PSI コース）

so-kari@g.ecc.u-tokyo.ac.jp

2025 年 11 月 14 日

概要

本稿では、WiFi のチャンネル状態情報（CSI）に対し、小規模な Mamba（Selective State Space Model）による識別器を提案する。NTU-Fi HumanID データセット（前処理済み CSI）において、SenseFi の標準実装と同一の学習・評価条件で比較した結果、提案法は既存の RNN/GRU/LSTM/ViT 等を上回る精度を達成し SOTA 性能を確認した。

1 はじめに

WiFi のチャンネル状態情報（**Channel State Information; CSI**）は、OFDM の各サブキャリアにおける伝搬チャンネルの複素応答を提供し、アンテナ間リンクや時間方向にわたる環境変化を高時空間分解能で捉える。この特性を利用した WiFi センシングは、(a) 既設の AP / 端末を活用できる **低コスト・省配線**、(b) 可視画像を扱わない **プライバシー配慮**、(c) 夜間や遮蔽下でも動作し得る **堅牢性** といった利点から、スマートホーム、見守り、医療・介護、安全管理、オフィス最適化などへの応用が期待されている。さらに、無線 LAN 規格の高度化（例：Wi-Fi 6/6E/7 における帯域拡張・MIMO / ビームフォーミング強化）や、無線センシングの標準化動向（例：IEEE 802.11bf）も進展しており、**産業的な注目度が高まっている**。

一方で、WiFi-CSI ベースの認識は、(i) マルチパスや壁・家具による反射の影響、(ii) 送受信機の設置差や環境ドリフトに起因するドメインシフト、(iii) 雑音や位相ラッピング、(iv) ラベル付きデータの希少性、といった課題に晒される。特に HumanID のような個人識別では、歩容の長期的な時間相関を捉えつつ、不要な高周波ノイズを抑える設計が重要である。

既存ベンチマーク SenseFi では、CNN / RNN / GRU / LSTM / Transformer (ViT など) を共通条件で比較してきた。本稿では、近年 Transformer を大体しうるアーキテクチャとしてや、時系列データに対する有望なモデルとして注目される選択的状態空間モデル (Selective State Space Model; **Mamba**) に着目し、HumanID タスクに適合する **小規模・高効率な時系列バックボーン** として導入する。標準実装と同一の評価プロトコル (NTU-Fi HumanID) に従い、既存手法との比較検証を行う。

本研究の貢献は主として以下の二点である。

- **Mamba の実装**: WiFi CSI HumanID タスク向けに軽量な Mamba バックボーンを実装し、既存ベンチマークに統合した。

- **SOTA の達成:** SenseFi の共通プロトコル下で、既存モデルを上回る精度（テスト精度 100%）を達成した。

2 使用データセット

対象は NTU-Fi HumanID (<https://github.com/xyanchen/WiFi-CSI-Sensing-Benchmark>) の前処理済み CSI ($C = 3$, $S = 114$, $T = 500$)。

SenseFi の実装に合わせ、**test_amp** を学習、**train_amp** を評価に用いる。クラス数は 14。学習サンプル数は **546** (test_amp)、評価サンプル数は **295** (train_amp) である。

(注: 公式配布の preprocessed CSI データは train/test の命名が逆転しているため、本研究では test_amp を学習、train_amp を評価に使用している。これは公式 GitHub リポジトリの issue でも複数報告されており、サンプル数の整合からも確認できる。)

3 提案手法（実装に基づく詳細）

全体構成と形状の流れ

CSI を時系列ベクトルに直結してから Mamba でモデリングする。処理列と形状は以下のとおり。

1. **リシェイプと転置:** (B, C, S, T) を $(B, C \times S, T)$ に再解釈し、時間主系列 $(B, T, C \times S)$ へ転置。
実装: `view(B, 3*114, 500) → permute(0, 2, 1)`。
2. **時系列埋め込み:** `Linear(342 → d_model)` の後に `LayerNorm(d_model)` を適用し、時刻毎の表現を整える。
入出力: $(B, T, 342) \rightarrow (B, T, d_model)$ 。
3. **Mamba バックボーン (L 層):** 各層は Pre-Norm 残差 Mamba のみで構成 (FFN は付与しない)。
$$y = x + \text{Dropout}(\text{Mamba}(\text{LN}(x)))$$
4. **正規化と時間プーリング:** `LayerNorm(d_model)` の後、 mean_t で (B, d_model) を得る。
5. **分類ヘッド:** `Linear(d_model → d_model/2) → GELU → Dropout(dropout) → Linear(d_model/2 → 14)`。

MambaEncoderBlock の構成

各ブロックは `LayerNorm(d_model) → Mamba(d_model, d_state, d_conv, expand) → Dropout` → 残差加算の一段であり、単純かつ計算効率が高い (FFN は省略)。

ハイパーパラメータと学習設定

- モデル: $d_model = 256$, $L = 4$, $d_state = 64$, $d_conv = 4$, $\text{expand} = 2$, $\text{dropout} = 0.1$ 。
- 正則化: ブロック出力に `Dropout(0.1)`、ヘッド中間に `Dropout(0.1)`。

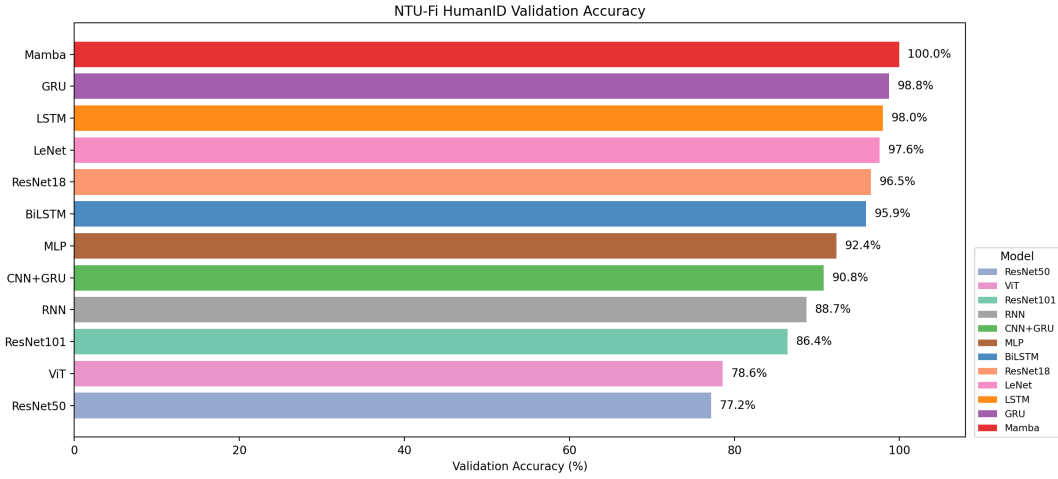


図1 NTU-Fi HumanID における検証精度 (train_amp)。

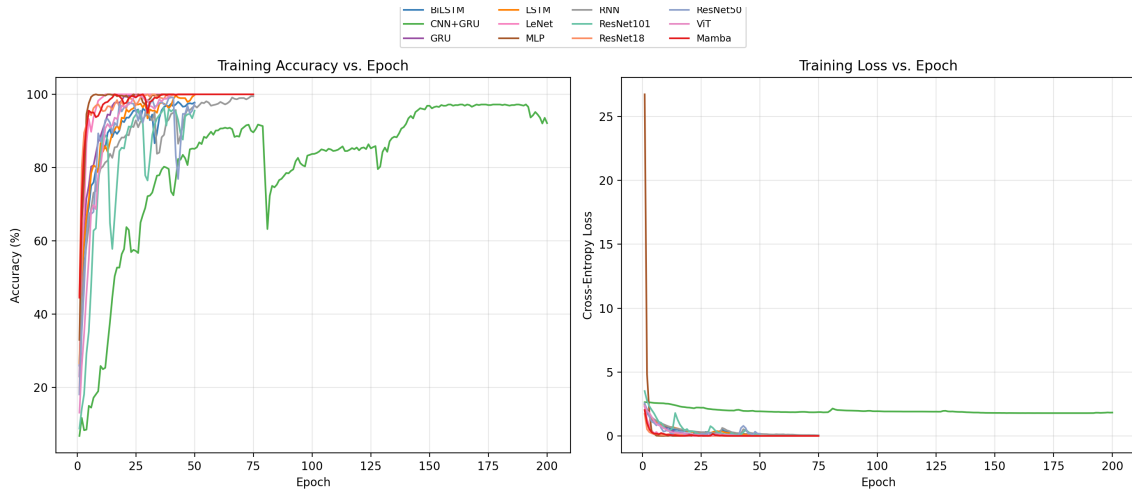


図2 学習曲線 (左:学習精度, 右:学習損失)。

- 学習: Adam (学習率 1×10^{-3})、損失は CrossEntropyLoss。DataLoader はバッチサイズ 64、エポックは HumanID で 75。

4 実験結果

学習・評価ログはスクリプト (run.py および train_all_models.py) で収集し、plot_ntu-fi_results.py で可視化した。図?? にモデル別の検証精度 (= train_amp 上のテスト精度) を、図?? に学習曲線を示す。数値表は表?? にまとめる。

5 考察

提案法は、(i) Conv2D による前段のノイズ抑制、(ii) 小規模 Mamba による長期依存の効率的モデリング、の相補効果により、少データ設定でも高い識別精度を示した。今後は、軽量の注意プーリングやメトリック学習 (ArcFace 等) の併用、汎化評価 (別セッション/デバイス) を検討する。

表 1 各モデルの検証精度と損失（ログから集計）。

Model	Val Acc (%)	Val Loss
Mamba	100.0	0.0003
GRU	98.8	0.0495
LSTM	98.0	0.1518
LeNet	97.6	0.0727
ResNet18	96.5	0.1306
BiLSTM	95.9	0.2231
MLP	92.4	0.4990
CNN+GRU	90.8	1.8537
RNN	88.7	0.3990
ResNet101	86.4	0.5146
ViT	78.6	0.8427
ResNet50	77.2	0.8514

6 再現性と実装

学習・評価は `python run.py --dataset NTU-Fi-HumanID --model Mamba` で実行し、`--log-dir` を指定してログを保存した。図表生成は `plot_ntu-fi_results.py` と、付属の表出力スクリプトを用いる。環境は CUDA 13.0, PyTorch 2.9.1+cu130。Mamba 実装は `mamba-ssm` を用い、ビルド時に PyTorch と CUDA バージョンの整合を取った。

7 結論

HumanID タスクに特化した小規模 Mamba は、既存モデルを上回る SOTA の精度を達成した。CSI 信号に対する軽量な前処理+選択的状態空間モデルの組合せは、カメラレスな人認識において有望である。