MedNeXt Model Evaluation Report

Liver Fibrosis Staging (F0–F4)

ConvNeXt-Tiny Backbone · 5-Class Classification

⚡ Performance Summary

Best Validation Accuracy: 98.66%

Test Accuracy: 98.66%

Cohen's Kappa: 0.9824 | F1-Score (Macro): 0.9803

Training Configuration

Backbone: ConvNeXt Tiny (timm) | Epochs: 50 | Batch Size: 16

Optimizer: AdamW | LR: 0.0001 | Scheduler: CosineAnnealingLR

Label Smoothing: 0.1 | Class Balancing: WeightedRandomSampler

Train Samples: 5058 | Val Samples: 1265

Detailed Performance Metrics
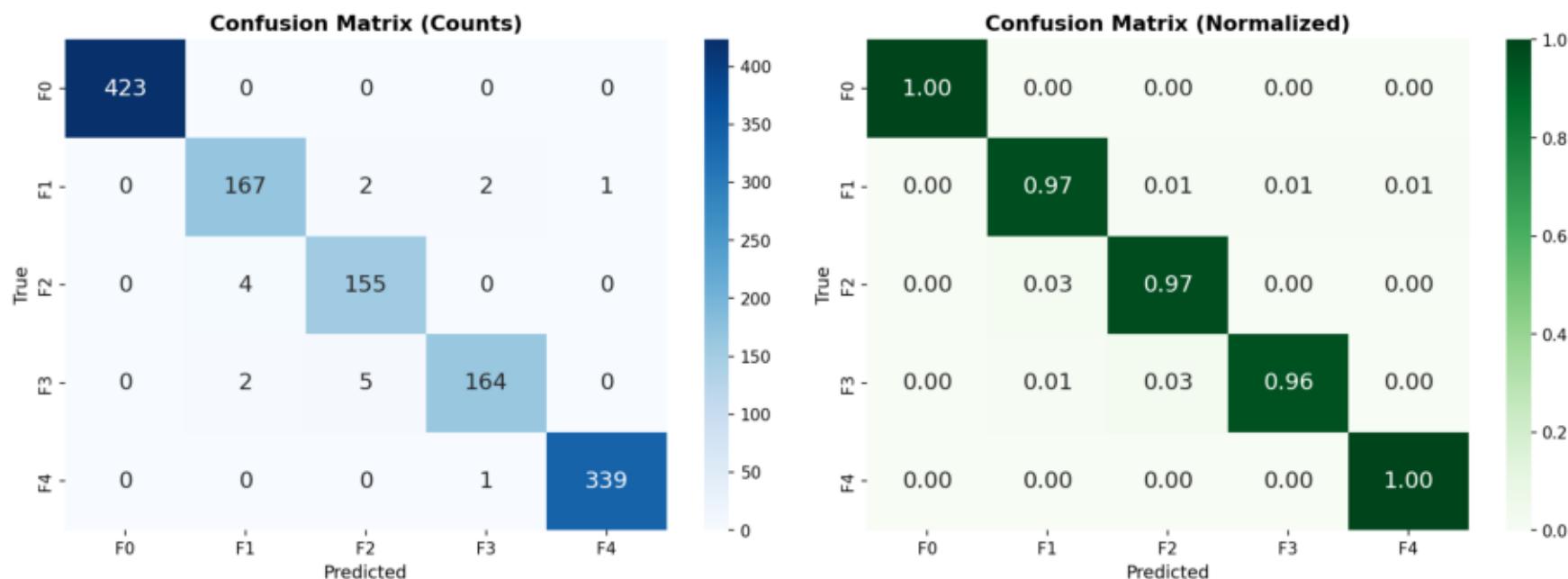
Metric	Value
Accuracy	98.66%
Cohen's Kappa	0.9824
Precision (Macro)	0.9802
Recall (Macro)	0.9804
F1-Score (Macro)	0.9803

Per-Class Performance
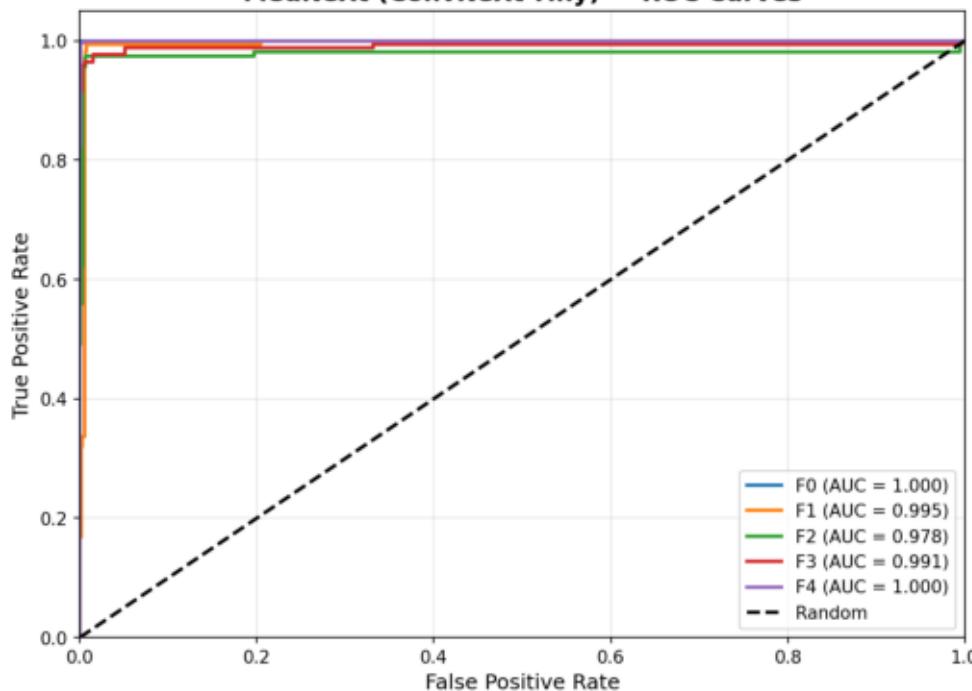
Class	Precision	Recall	F1-Score	AUC	Support
F0	1.0000	1.0000	1.0000	1.0000	423
F1	0.9653	0.9709	0.9681	0.9948	172
F2	0.9568	0.9748	0.9657	0.9784	159
F3	0.9820	0.9591	0.9704	0.9913	171
F4	0.9971	0.9971	0.9971	0.9998	340

MedNeXt Visualizations

MedNeXt (ConvNeXt-Tiny) — Confusion Matrices

Confusion Matrix (Counts)

Confusion Matrix (Normalized)
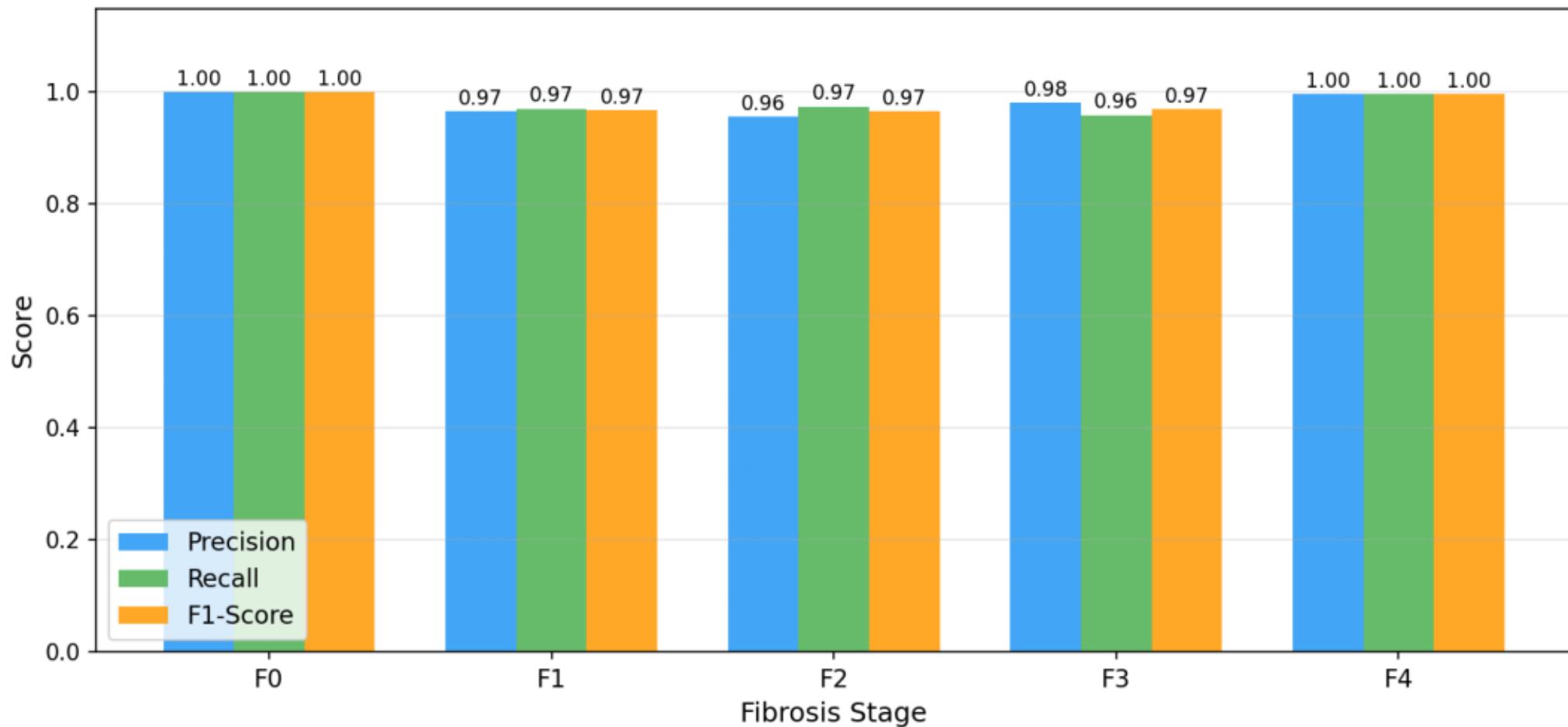
MedNeXt (ConvNeXt-Tiny) — ROC Curves

- F0 (AUC = 1.000)
- F1 (AUC = 0.995)
- F2 (AUC = 0.978)
- F3 (AUC = 0.991)
- F4 (AUC = 1.000)
- Random

Per-Class Performance Breakdown

MedNeXt — Per-Class Performance

Classification Report

	precision	recall	f1-score	support
F0	1.00	1.00	1.00	423
F1	0.97	0.97	0.97	172
F2	0.96	0.97	0.97	159
F3	0.98	0.96	0.97	171
F4	1.00	1.00	1.00	340
accuracy			0.99	1265
macro avg	0.98	0.98	0.98	1265
weighted avg	0.99	0.99	0.99	1265