Comparative Analysis of Deep Learning Models

for Liver Fibrosis Staging

Research Report

Top Performing Model

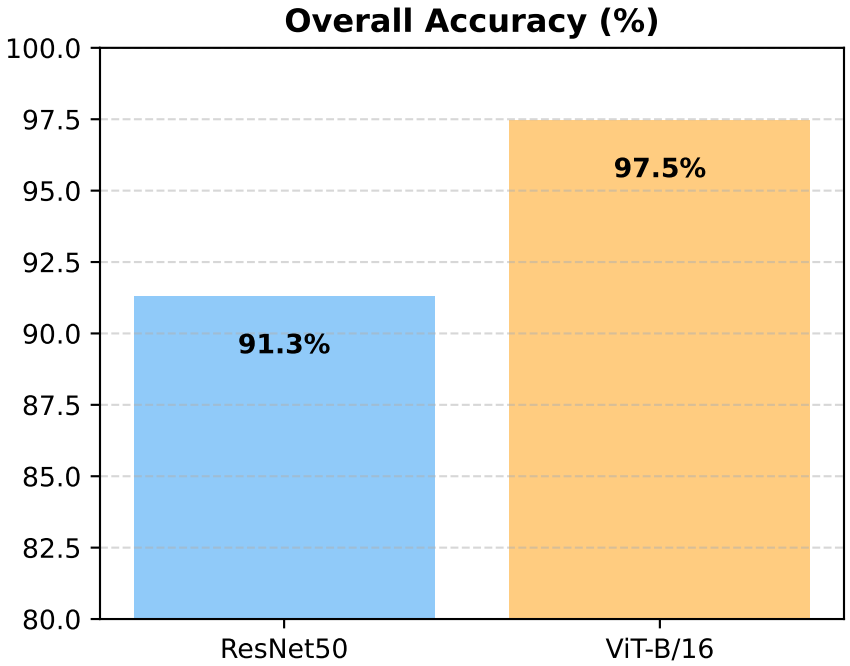
VIT-B/16

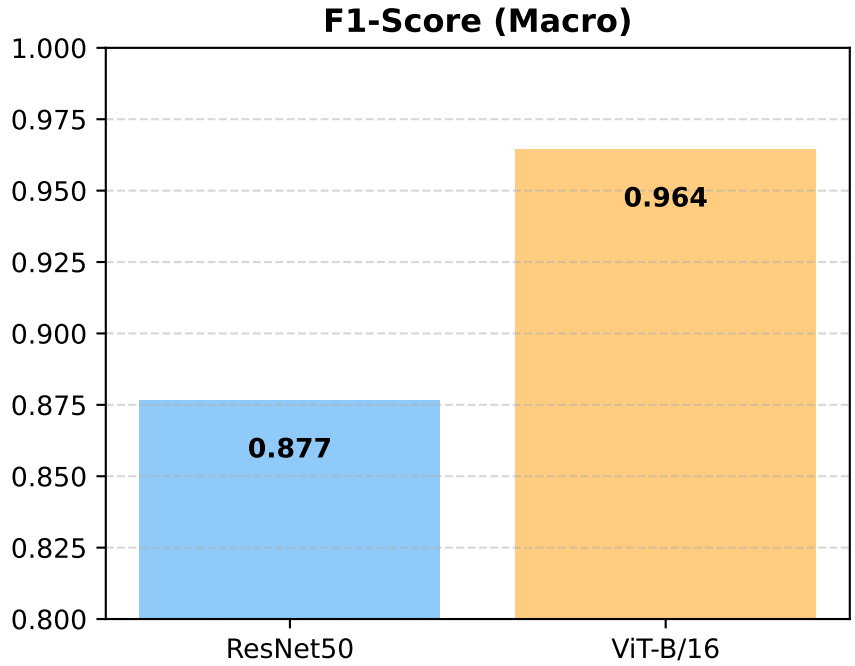
Accuracy: 97.47%

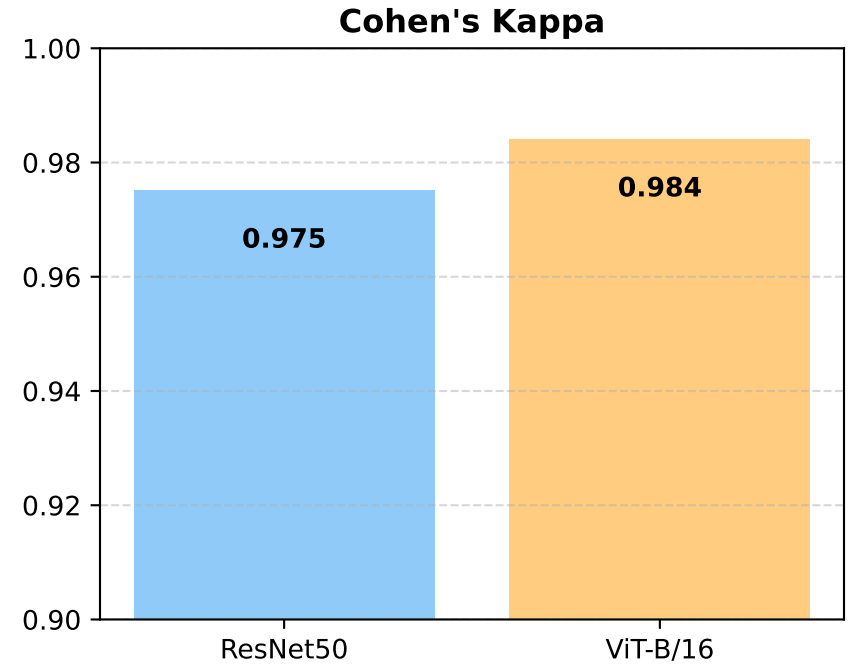
Executive Summary

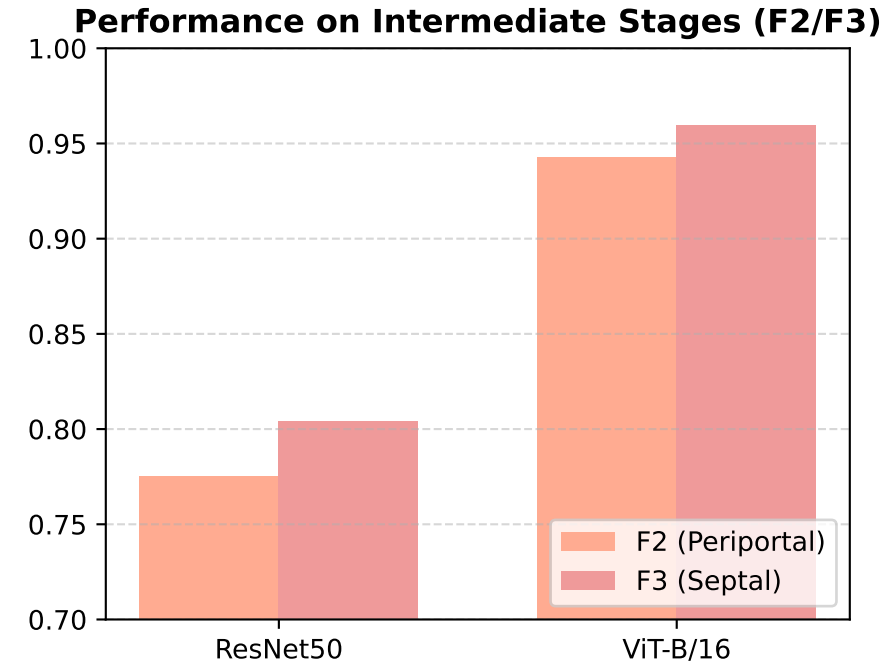
This study evaluates the performance of three distinct deep learning architectures for the automated staging of liver fibrosis from histopathology images: ResNet50 (baseline CNN), EfficientNet-V2 (optimized CNN), and Vision Transformer (ViT-B/16). Key Findings: 1. Superiority of Transformers: The Vision Transformer (ViT-B/16) achieved the highest overall accuracy (97.47%), outperforming both CNN-based approaches. This suggests that the self-attention mechanism is highly effective at capturing global tissue patterns indicative of fibrosis. 2. Resilience in Intermediate Stages: A critical challenge in fibrosis staging is distinguishing between intermediate stages (F2, F3). The ViT model demonstrated significantly higher sensitivity and precision for these classes compared to the ResNet50 baseline. 3. Efficiency vs. Performance: EfficientNet-V2 provided a very competitive performance (96.60%) with a lighter computational footprint, making it a viable alternative for resource-constrained deployments. 4. Clinical Relevance: The high Cohen's Kappa scores (>0.98 for top models) indicate excellent agreement with ground truth, supporting the potential utility of these models as decision support tools in clinical pathology workflows.

Model Performance Comparison







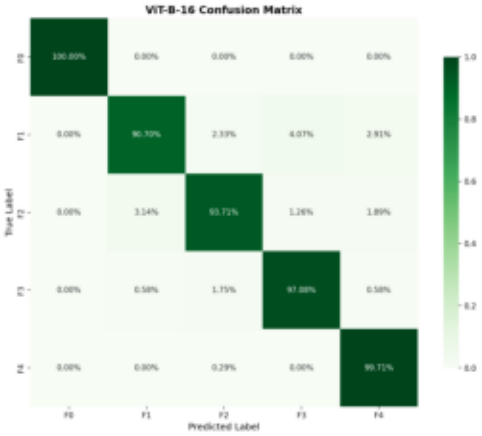


Detailed Performance Metrics

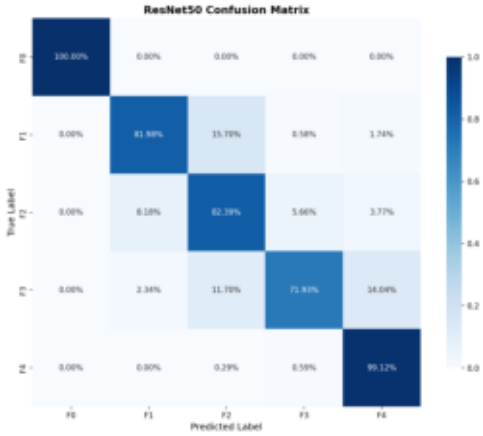
Metric	RESNet	VIT
Accuracy	91.30%	97.47%
Precision (Macro)	0.8892	0.9669
Recall (Macro)	0.8708	0.9624
F1-Score (Macro)	0.8766	0.9644
Cohen's Kappa	0.9751	0.9841
ROC AUC (Macro)	0.9889	0.9981
F1 - Stage F0	1.0000	1.0000
F1 - Stage F1	0.8545	0.9341
F1 - Stage F2	0.7751	0.9430
F1 - Stage F3	0.8039	0.9595
F1 - Stage F4	0.9493	0.9855

Confusion Matrices

ViT-B/16

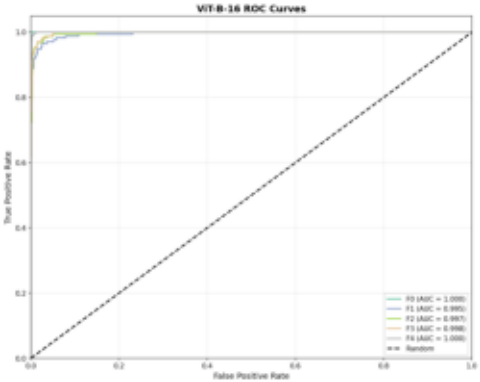


ResNet50



ROC Curves

ViT-B/16



ResNet50