

Ensemble Evaluation Summary

4-model ensemble for bone fracture classification

Rank	Model	Accuracy	Precision	Recall	F1	AUC
1	Ensemble (Soft Voting)	0.99935	0.99940	0.99935	0.99937	1.00000
2	Swin Transformer (Ti	0.99903	0.99905	0.99900	0.99902	1.00000
3	MaxViT-Tiny	0.99805	0.99815	0.99799	0.99807	0.99999
4	ConvNeXt V2 Base	0.99578	0.99579	0.99581	0.99580	0.99989
5	EfficientNetV2-S	0.99578	0.99592	0.99574	0.99583	0.99995

Notes:

- Metrics computed on full test set (3082 images).
- Ensemble uses soft voting over model probabilities.
- See results/ensemble_eval for confusion matrices and ROC curves.