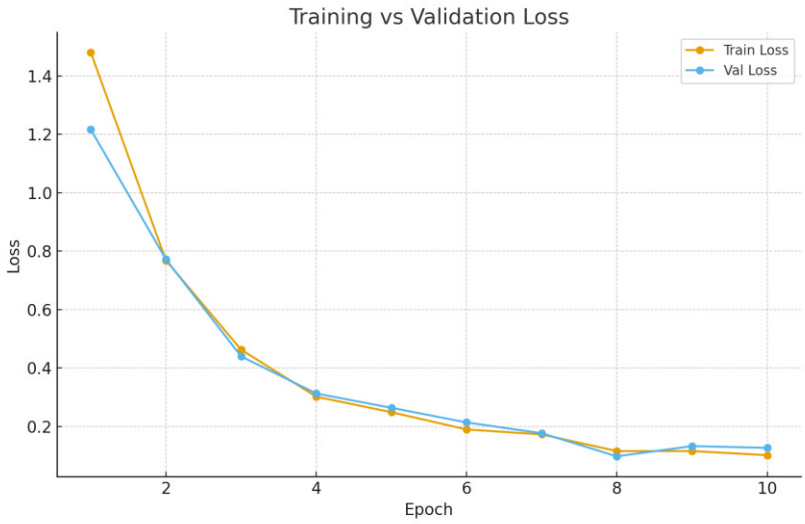
Klasifikacija instrumenata – izveštaj

What We Changed

Training was extended to 10 epochs, and instead of uniform sampling (K=800), we used the WeightedRandomSampler for class balancing.

Training Trends

Ep	Train Loss	Train Acc	Val Loss	Val Acc
1	1.480	0.439	1.216	0.542
2	0.768	0.700	0.772	0.732
3	0.463	0.824	0.440	0.875
4	0.302	0.890	0.313	0.915
5	0.249	0.911	0.264	0.940
6	0.190	0.936	0.214	0.953
7	0.173	0.939	0.177	0.972
8	0.116	0.964	0.098	0.981
9	0.116	0.963	0.133	0.972
10	0.102	0.965	0.127	0.971



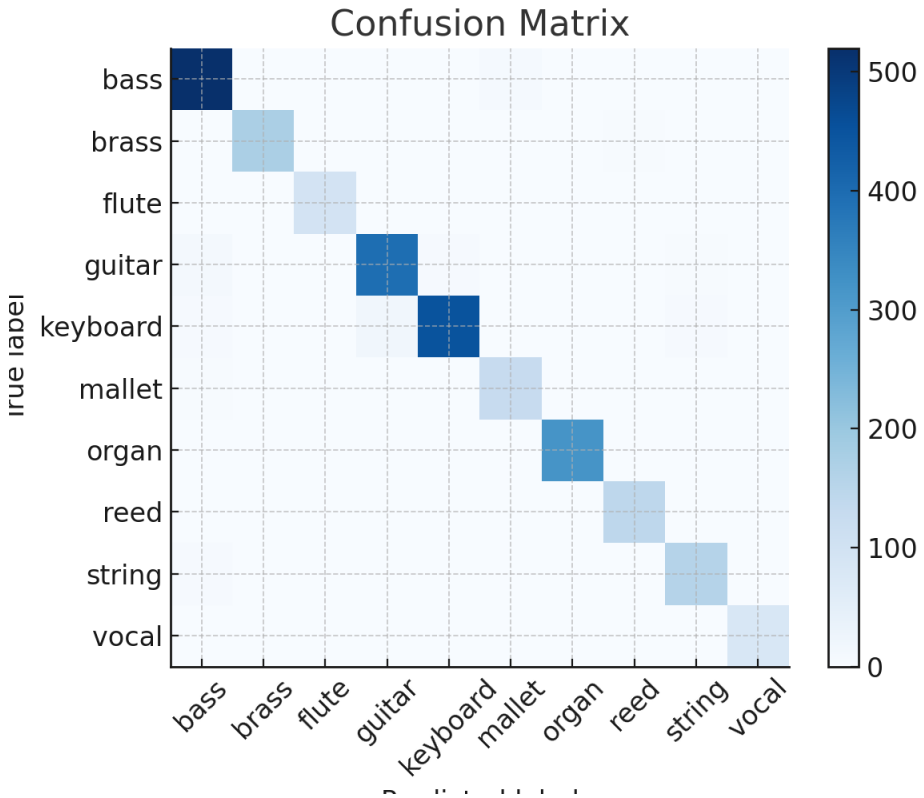
Performance per Class (Validation)

Class	Precision	Recall	F1-score	Support
bass	0.95	0.98	0.97	528
brass	1.00	0.98	0.99	177
flute	1.00	1.00	1.00	94
guitar	0.96	0.95	0.95	416
keyboard	0.98	0.94	0.96	481
mallet	0.94	0.96	0.95	132
organ	1.00	0.99	1.00	320
reed	0.97	1.00	0.98	144
string	0.95	0.97	0.96	163
vocal	1.00	1.00	1.00	81

Accuracy: 0.97 (N=2536)

Macro avg: Precision=0.97, Recall=0.98, F1=0.98

Weighted avg: Precision=0.97, Recall=0.97, F1=0.97



Conclusion

The model achieved a validation accuracy of 97% with very high F1-score values across all classes. The best performances were observed in the flute and vocal classes (F1=1.00). Minor confusion occurs between the guitar and keyboard classes, but the overall error rate is minimal.