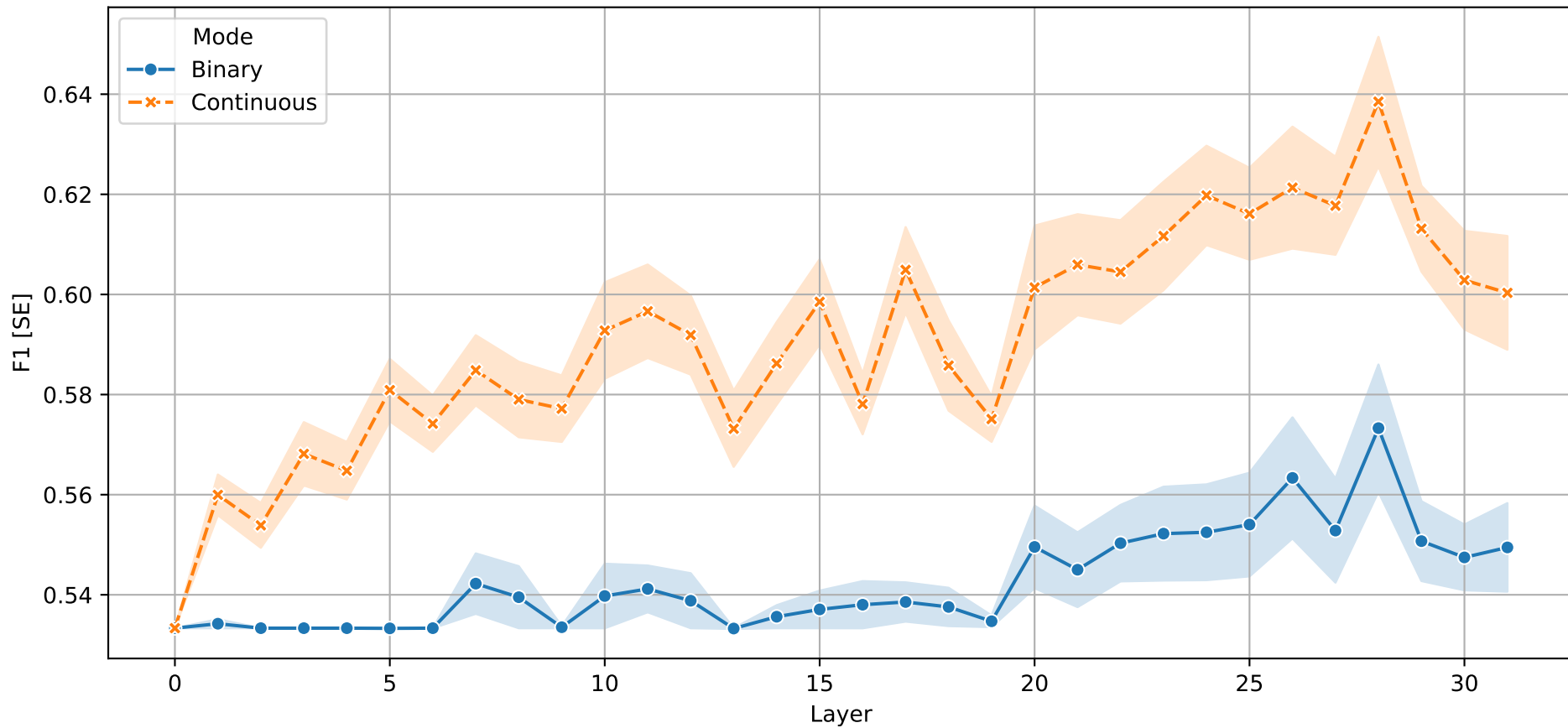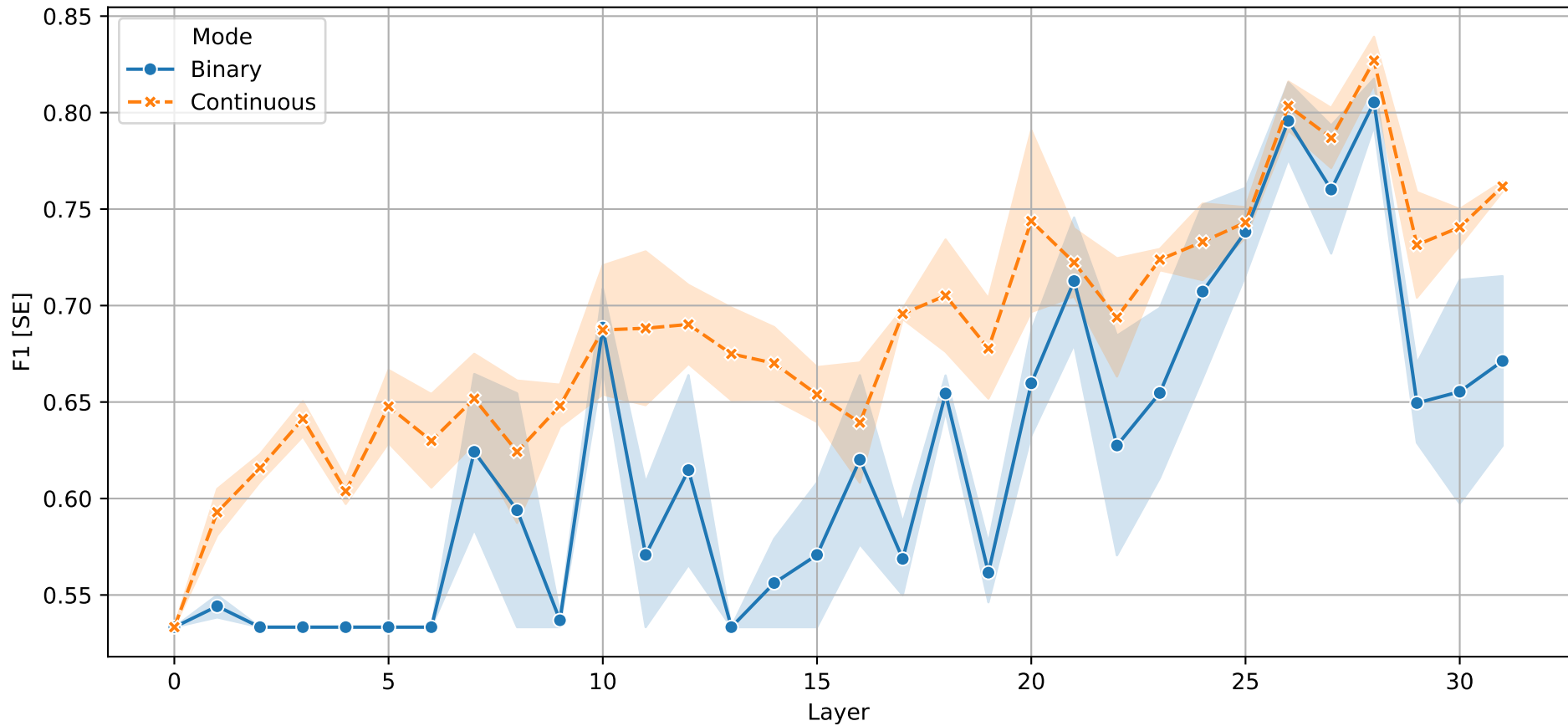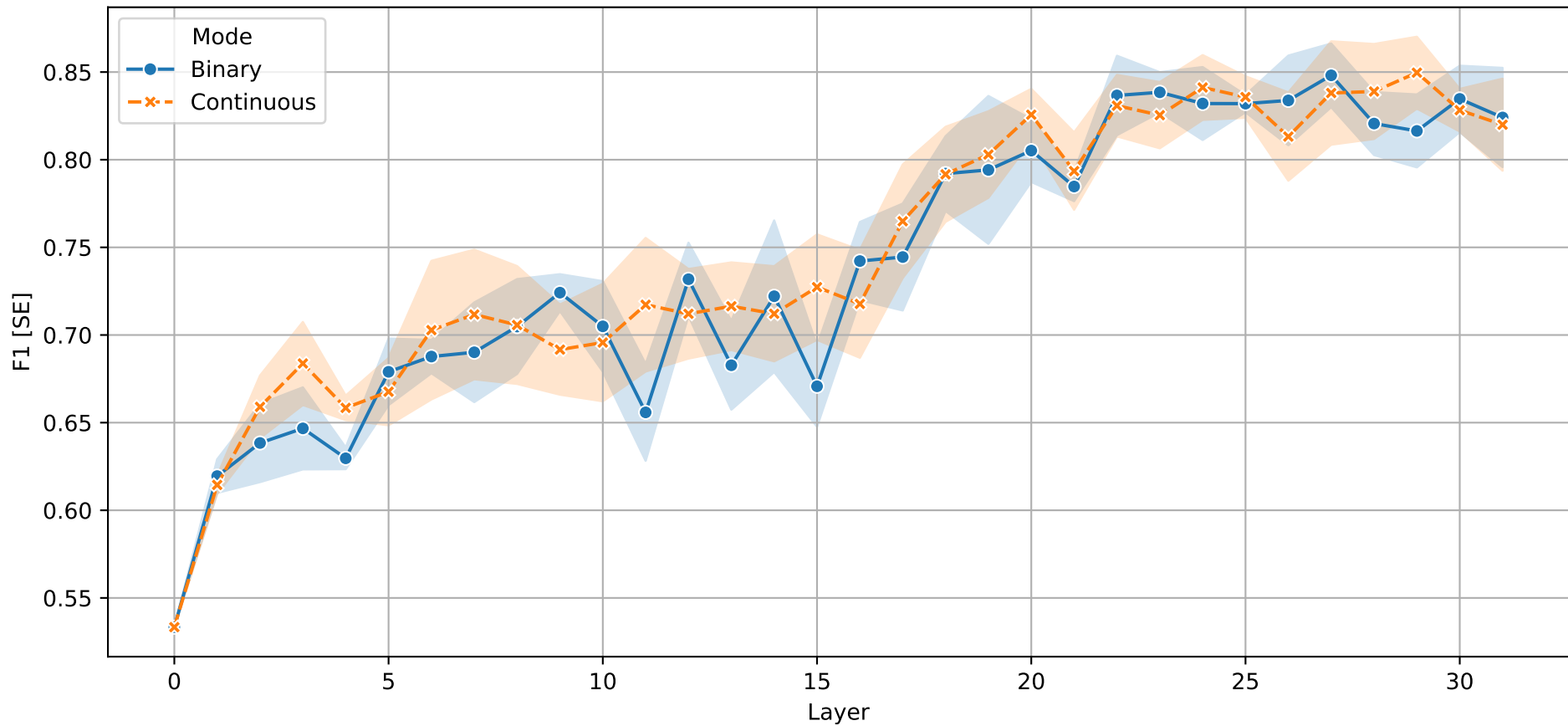F1 per Layer – Single Neuron Probing for landmarks
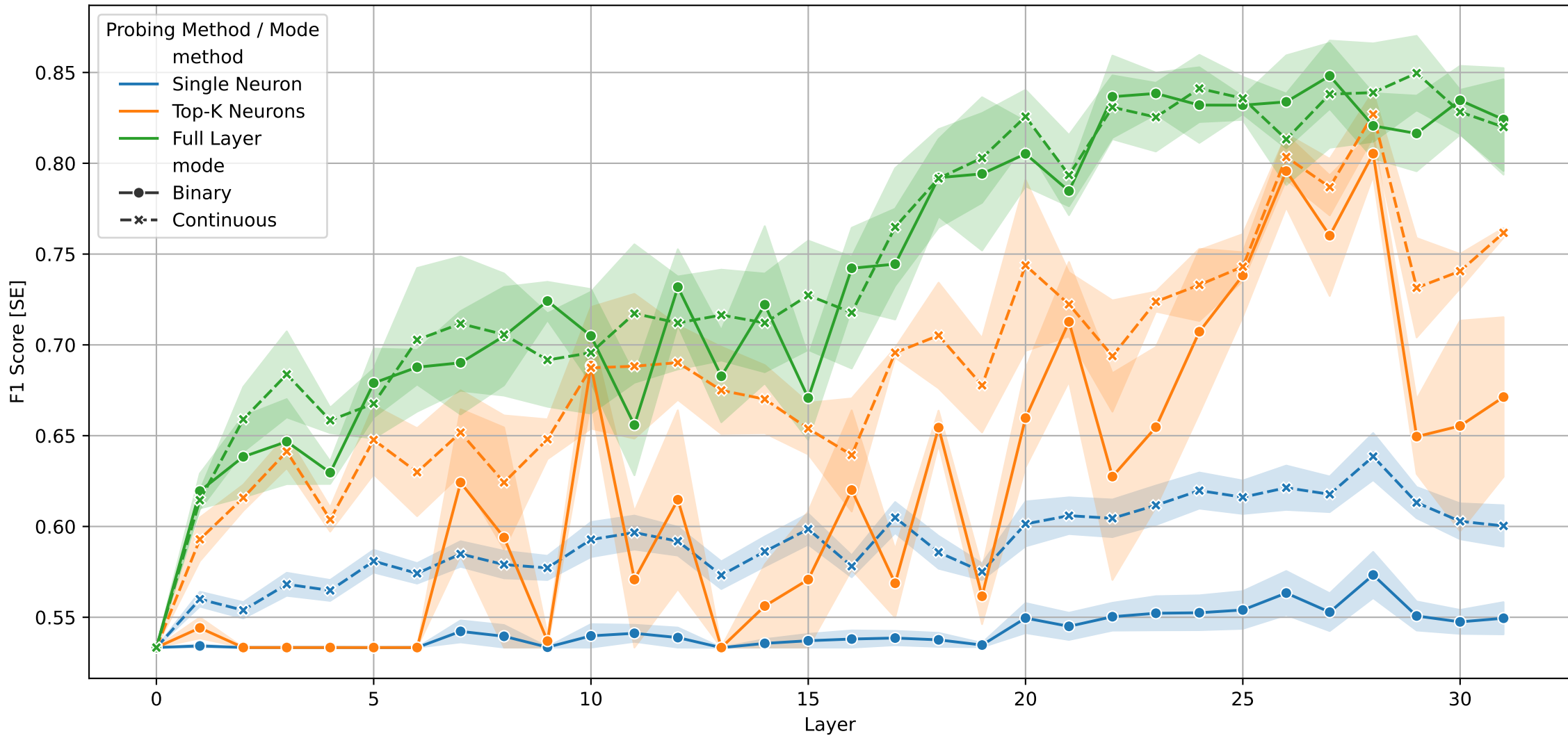
F1 per Layer – Top-K Neurons Probing for landmarks

F1 per Layer – Full Layer Probing for landmarks
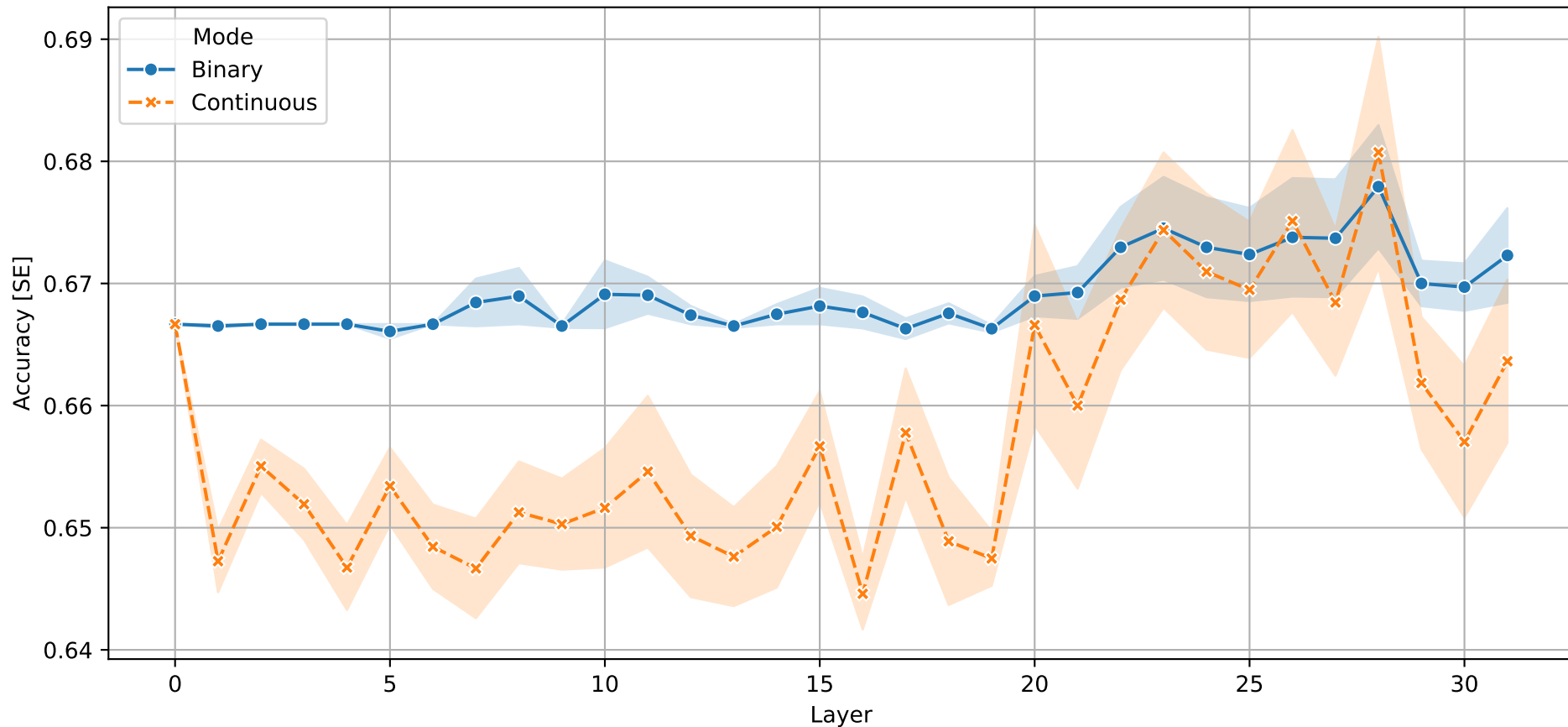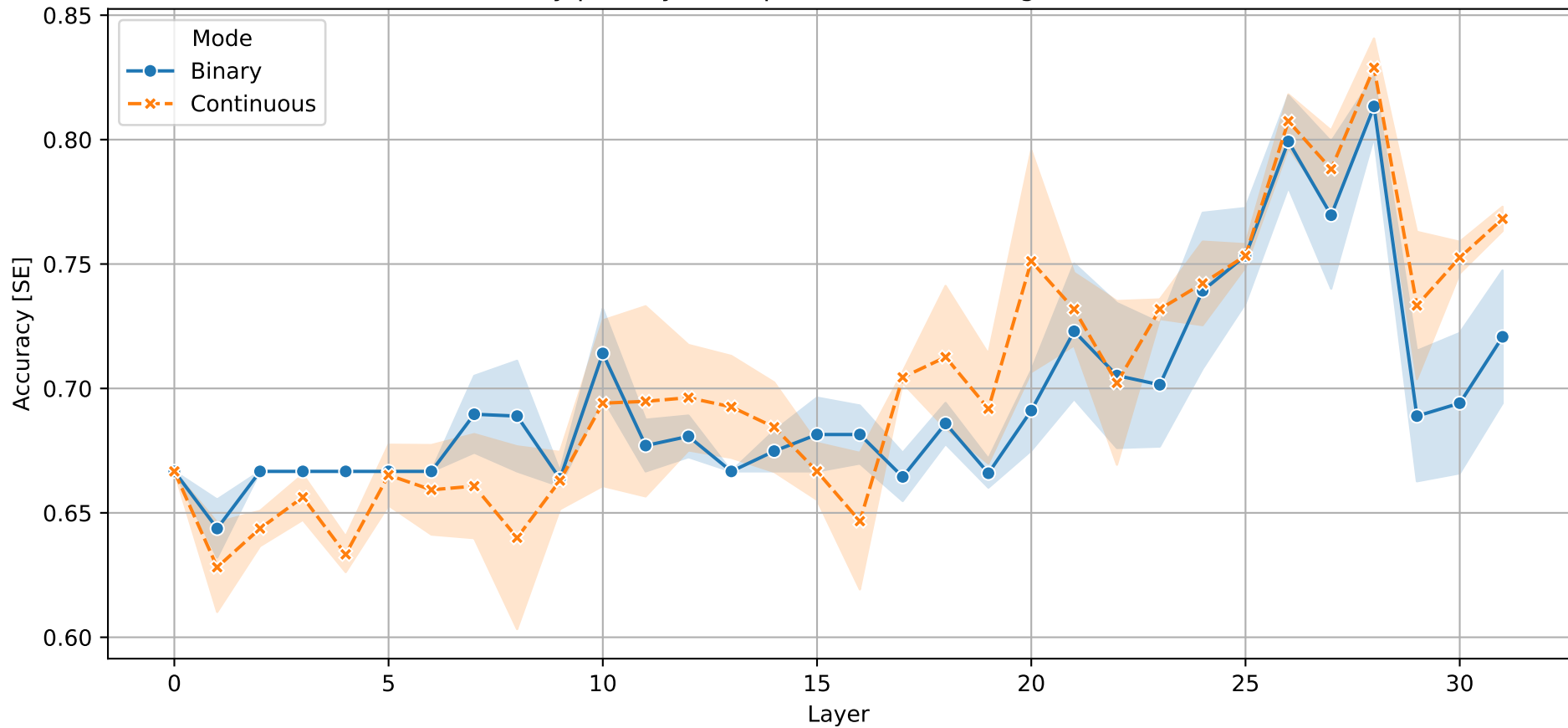
Overall F1 per Layer – All Methods for landmarks

F1 Score Summary by Probing Method

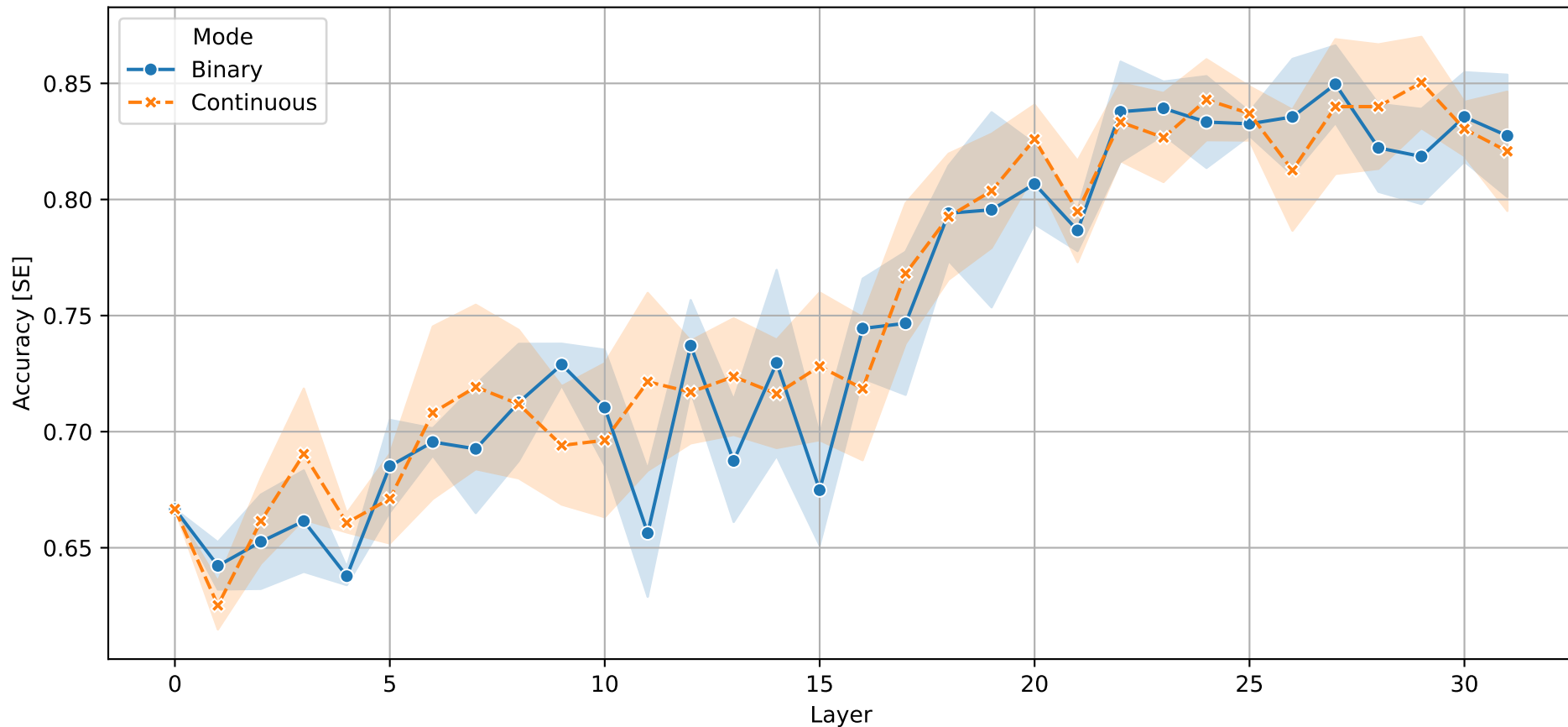| Method	Metric	Binary	Continuous
Full Layer	f1_best_layer	27.0	29.0
Full Layer	f1_max	0.8842	0.8967
Full Layer	f1_mean	0.7407	0.7477
Full Layer	f1_std	0.0879	0.0858
Single Neuron	f1_best_layer	28.0	28.0
Single Neuron	f1_max	0.7987	0.8431
Single Neuron	f1_mean	0.5426	0.591
Single Neuron	f1_std	0.0358	0.0512
Top-K Neurons	f1_best_layer	28.0	28.0
Top-K Neurons	f1_max	0.8353	0.8463
Top-K Neurons	f1_mean	0.6211	0.687
Top-K Neurons	f1_std	0.0926	0.0706

Accuracy per Layer – Single Neuron Probing for landmarks

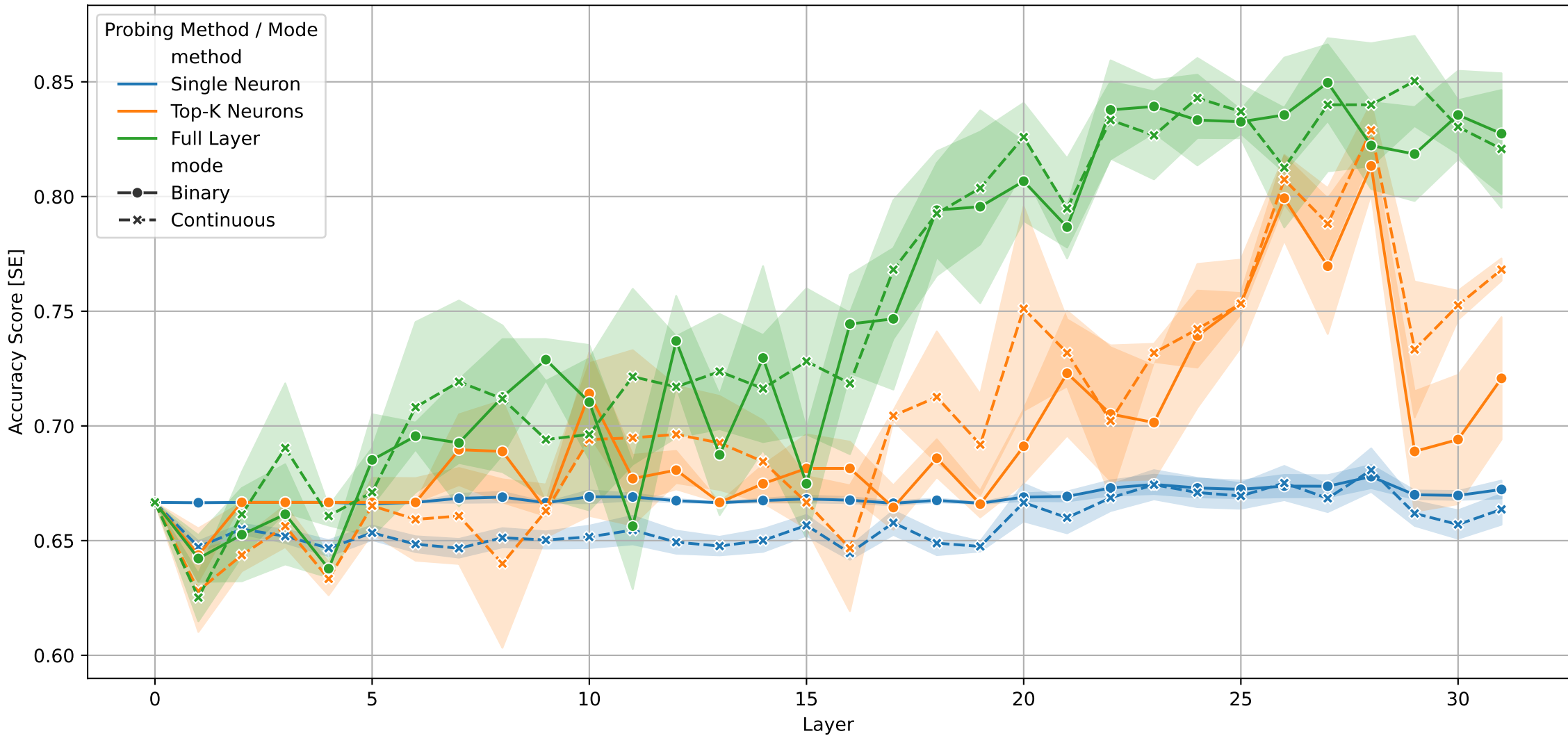Accuracy per Layer – Top-K Neurons Probing for landmarks

Accuracy per Layer – Full Layer Probing for landmarks

Overall Accuracy per Layer – All Methods for landmarks

Accuracy Score Summary by Probing Method

Method	Metric	Binary	Continuous
Full Layer	accuracy_best_layer	27.0	29.0
Full Layer	accuracy_max	0.8844	0.8978
Full Layer	accuracy_mean	0.7493	0.7547
Full Layer	accuracy_std	0.0775	0.0765
Single Neuron	accuracy_best_layer	28.0	28.0
Single Neuron	accuracy_max	0.8044	0.8467
Single Neuron	accuracy_mean	0.6692	0.6576
Single Neuron	accuracy_std	0.0134	0.0298
Top-K Neurons	accuracy_best_layer	28.0	28.0
Top-K Neurons	accuracy_max	0.8356	0.8467
Top-K Neurons	accuracy_mean	0.6962	0.7029
Top-K Neurons	accuracy_std	0.0467	0.0592