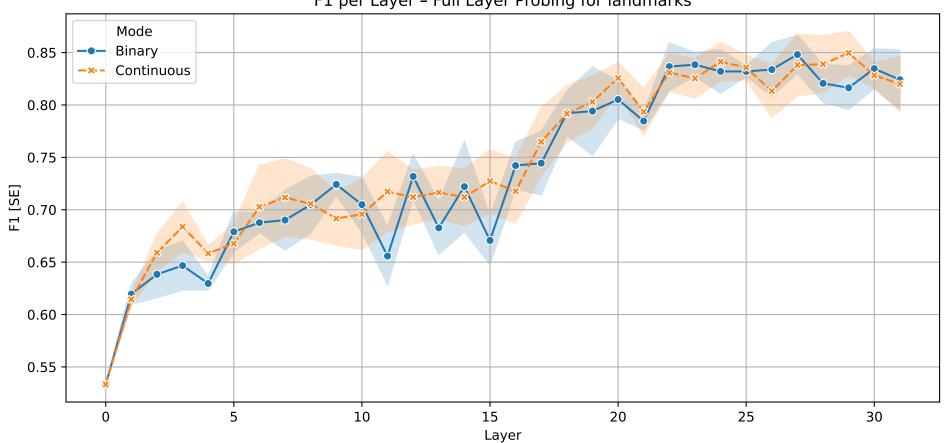
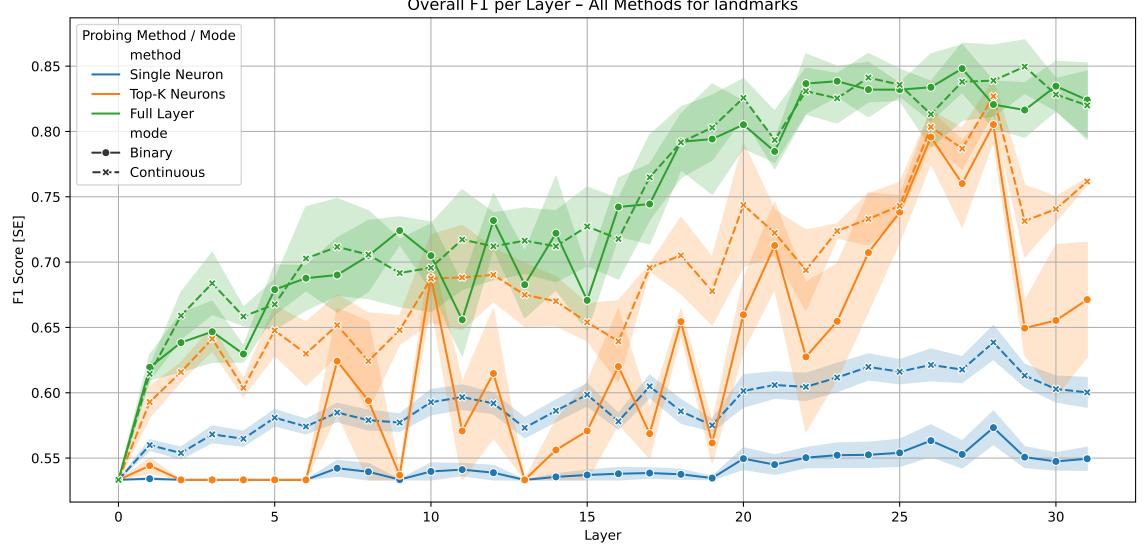
F1 per Layer - Single Neuron Probing for landmarks Mode Binary 0.64 -Continuous 0.62 F1 [SE] 0.58 0.56 0.54 20 25 30 10 15 Layer

F1 per Layer - Top-K Neurons Probing for landmarks 0.85 Mode Binary Continuous 0.80 0.75 <u>S</u> 0.70 0.65 0.60 0.55 10 20 15 25 30 Layer

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 f1 best layer

f1 may

f1 best layer

f1 max

f1 mean

f1 std

Full Layer

Full Layer

Top-K Neurons

Top-K Neurons

Top-K Neurons

Top-K Neurons

I uli Layei	TI_IIIdX	0.0042	0.0307
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

27.0

0.8842

28.0

0.8353

0.6211

0.0926

Continuous

29.0

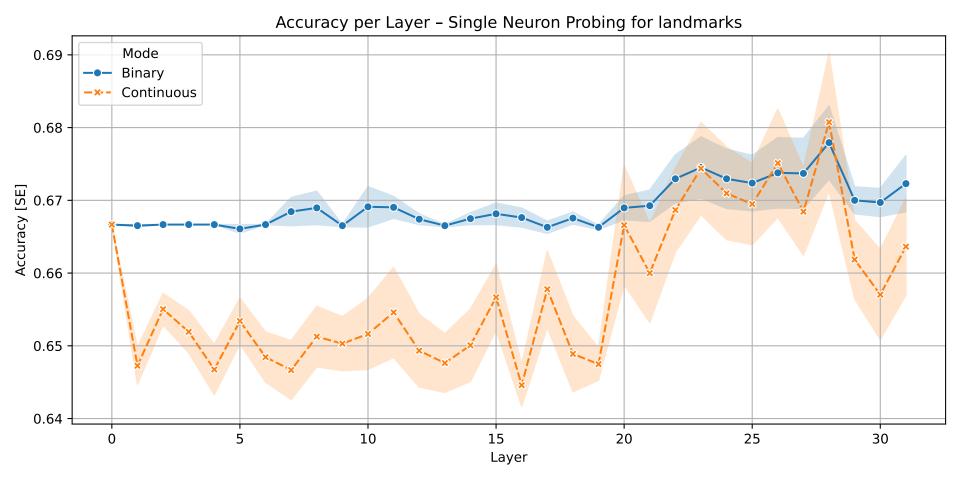
0.8067

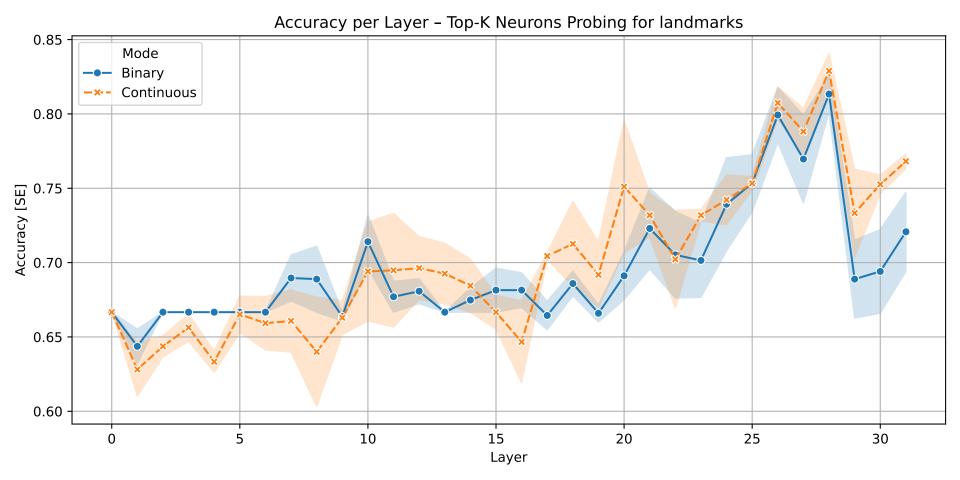
28.0

0.8463

0.687

0.0706





Accuracy per Layer - Full Layer Probing for landmarks Mode Binary 0.85 -Continuous 0.80 Accuracy [SE] 0.70 0.65 20 25 30 10 15 Layer

Overall Accuracy per Layer – All Methods for landmarks Probing Method / Mode method 0.85 Single Neuron Top-K Neurons Full Layer mode Binary 0.80 --x-- Continuous Accuracy Score [SE] 0.65 0.60 10 15 20 25 30 Layer

	Accuracy Score Summi	ary by ribbing Method
Method	Metric	Binary

accuracy best layer

accuracy max

accuracy max

accuracy mean

accuracy std

Full Layer

Full Laver

Top-K Neurons

Top-K Neurons

Top-K Neurons

· -y	1		
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

Accuracy Score Summary by Probing Method

27.0

0.8844

0.8356

0.6962

0.0467

Continuous

29.0

0.8978

0.8467

0.7029

0.0592