

Test Experiment Results:

Below we present the results of different methods of smoothing against not smoothing. Except for the scissors class in the P026_tissue2 video, smoothing improves every metric in every video. Following the results of P026_tissue2, we address why we suspect that there smoothing fails to improve the results.

Before delving into the results, we will reiterate the method of weighting:

$$\text{for } i \in \{1, \dots, |window\ size|\},$$

$$w_i = \frac{f(i+1)}{\sum_{j=1}^{|window\ size|} f(j+1)} \text{ s. t. } f(x) = \ln(x), \exp(x), x + \log(x), x$$

For example,

$$f(x) = x \text{ and } |window\ size| = 25 \text{ then } w_{25} = \frac{26}{325} = 0.11, w_1 = \frac{2}{325} = 0.006$$

$$f(x) = \log(x) \text{ and } |window\ size| = 25 \text{ then } w_{25} = \frac{\log(26)}{54.8} = 0.06, w_1 = \frac{\log(2)}{54.8} = 0.01$$

We will note as stated in our report that using the log transformation shifts importance away for newer frames and slows the transition between predictions.

The best performing methods are **bolded** and noticeable results are underlined.

P022 Balloon 1:

P022 Balloon1	Smoothing method	F1-Macro	Accuracy
	None	0.92	0.96
	Log	0.93	0.96
	Linear	0.94	0.965
	Log + Linear	0.94	0.965

P022 Balloon1	Smoothing method	Recall			
		Empty	Needle Driver	Forceps	Scissors
	None	0.92	0.99	0.97	0.85
	Log	0.93	0.99	0.98	0.92
	Linear	0.93	0.99	0.99	<u>0.93</u>
	Log + Linear	0.93	0.99	0.99	0.925

P022 Balloon1	Smoothing method	Precision			
		Empty	Needle Driver	Forceps	Scissors
	None	0.95	0.996	0.94	0.72
	Log	0.96	0.996	0.95	0.75
	Linear	0.97	0.996	0.95	0.76
	Log + Linear	0.97	0.996	0.95	0.76

P022 Balloon1	Smoothing method	F1			
		Empty	Needle Driver	Forceps	Scissors
	None	0.94	0.99	0.95	0.78
	Log	0.95	0.99	0.97	0.83
	Linear	0.95	0.99	0.97	0.84
	Log + Linear	0.95	0.99	0.97	0.84

P023 Tissue 2:

P023 Tissue2	Smoothing method	F1-Macro	Accuracy
	None	0.89	0.92
	Log	0.89	0.92
	Linear	0.90	0.93
	Log + Linear	0.90	0.93

P023 Tissue2	Smoothing method	Recall			
		Empty	Needle Driver	Forceps	Scissors
	None	0.88	0.99	0.78	0.93
	Log	0.88	0.996	0.80	0.92
	Linear	0.89	0.998	0.80	0.93
	Log + Linear	0.89	0.998	0.80	0.93

P023 Tissue2	Smoothing method	Precision			
		Empty	Needle Driver	Forceps	Scissors
	None	0.92	0.98	0.72	0.90
	Log	0.92	0.98	0.73	0.91
	Linear	0.93	0.98	0.74	0.92
	Log + Linear	0.93	0.98	0.745	0.92

P023 Tissue2	Smoothing method	F1			
		Empty	Needle Driver	Forceps	Scissors
	None	0.90	0.99	0.75	0.92
	Log	0.90	0.99	0.76	0.915
	Linear	0.91	0.99	0.77	0.92
	Log + Linear	0.91	0.99	0.77	0.92

P024 Balloon 1:

P024 Balloon1	Smoothing method	F1-Macro	Accuracy
	None	0.88	0.95
	Log	0.89	0.95
	Linear	0.91	0.96
	Log + Linear	0.91	0.96

P024 Balloon1	Smoothing method	Recall			
		Empty	Needle Driver	Forceps	Scissors
	None	0.95	0.96	-	0.79
	Log	0.96	0.97	-	0.81
	Linear	0.96	0.97	-	0.856
	Log + Linear	0.96	0.97	-	0.853

P024 Balloon1	Smoothing method	Precision			
		Empty	Needle Driver	Forceps	Scissors
	None	0.98	0.95	-	0.67
	Log	0.98	0.96	-	0.69
	Linear	0.98	0.96	-	0.726
	Log + Linear	0.98	0.96	-	0.728

P024 Balloon1	Smoothing method	F1			
		Empty	Needle Driver	Forceps	Scissors
	None	0.96	0.96	-	0.73
	Log	0.97	0.96	-	0.75
	Linear	0.97	0.97	-	0.785
	Log + Linear	0.97	0.97	-	0.786

We will note that forceps are not used in P024 Balloon 1.

P025 Tissue 2:

P025 Tissue2	Smoothing method	F1-Macro	Accuracy
	None	0.92	0.945
	Log	0.93	0.95
	Linear	0.93	0.95
	Log + Linear	0.93	0.95

P025 Tissue2	Smoothing method	Recall			
		Empty	Needle Driver	Forceps	Scissors
	None	0.91	0.98	0.93	0.93
	Log	0.91	0.98	0.93	0.94
	Linear	0.92	0.98	0.94	0.90
	Log + Linear	0.925	0.98	0.94	0.90

P025 Tissue2	Smoothing method	Precision			
		Empty	Needle Driver	Forceps	Scissors
	None	0.96	0.98	0.845	0.85
	Log	0.965	0.98	0.84	0.86
	Linear	0.97	0.98	0.865	0.87
	Log + Linear	0.97	0.98	0.865	0.88

P025 Tissue2	Smoothing method	F1			
		Empty	Needle Driver	Forceps	Scissors
	None	0.935	0.98	0.885	0.89
	Log	0.94	0.98	0.885	0.90
	Linear	0.95	0.98	0.90	0.89
	Log + Linear	0.95	0.98	0.90	0.89

P026 Tissue 2:

P026 Tissue2	Smoothing method	F1-Macro	Accuracy
	None	0.87	0.92
	Log	0.88	0.93
	Linear	0.88	0.93
	Log + Linear	0.87	0.93

P026 Tissue2	Smoothing method	Recall			
		Empty	Needle Driver	Forceps	Scissors
	None	0.85	0.99	0.865	<u>0.92</u>
	Log	0.87	0.99	0.875	<u>0.86</u>
	Linear	0.875	0.99	0.88	0.85
	Log + Linear	0.87	0.99	0.88	0.80

P026 Tissue2	Smoothing method	Precision			
		Empty	Needle Driver	Forceps	Scissors
	None	0.84	0.97	0.93	0.66
	Log	0.85	0.97	0.945	0.67
	Linear	0.86	0.97	0.95	0.67
	Log + Linear	0.86	0.97	0.95	0.66

P026 Tissue2	Smoothing method	F1			
		Empty	Needle Driver	Forceps	Scissors
	None	0.85	0.98	0.90	<u>0.77*</u>
	Log	0.86	0.98	0.91	<u>0.75</u>
	Linear	0.87	0.98	0.915	0.75
	Log + Linear	0.87	0.98	0.915	0.72

*We note that here the no smoothing method trumps the recall of the no smoothing method trump the smoothing methods. This is likely because in this video the surgeon holds the scissors for short periods of time. At first 9 frames then 23 frames then 198 frames.

In comparison, for P024_balloon1 the scissors are held for 237 frames then 251 frames then 291. Our window was 25 frames long and therefore unable to react quickly enough to these short windows.

A direction for future research is to find a combination of window size and smoothing method that responds quickly enough to improve recall on short periods of using a tool. At the time, this combination would need to preserve the quality found in our smoothing method of reducing quick switches caused by errors in the object detection model.