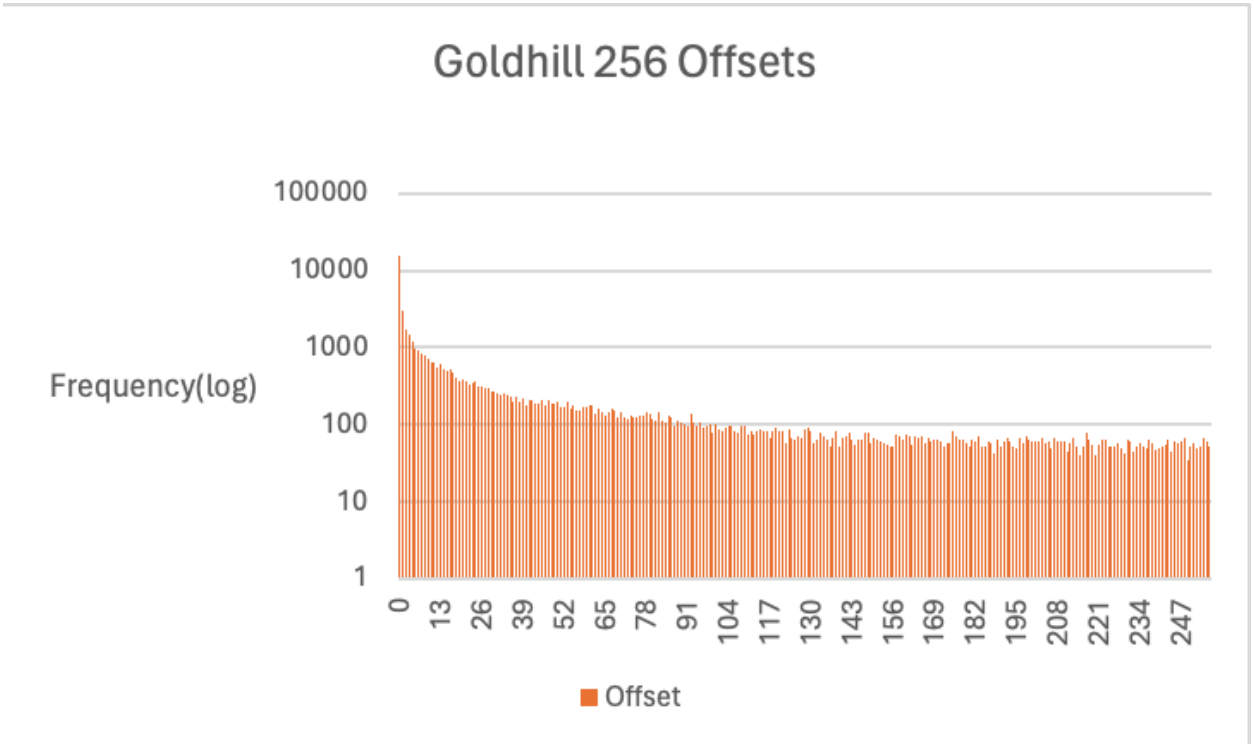
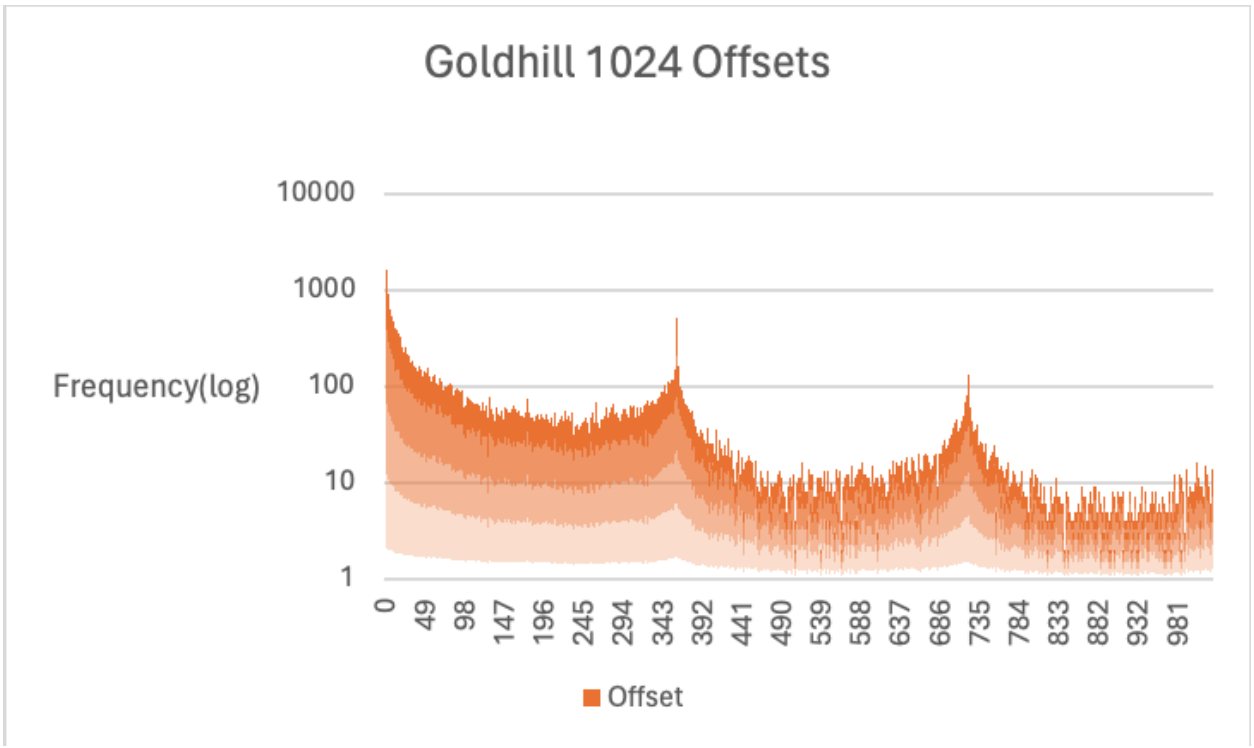
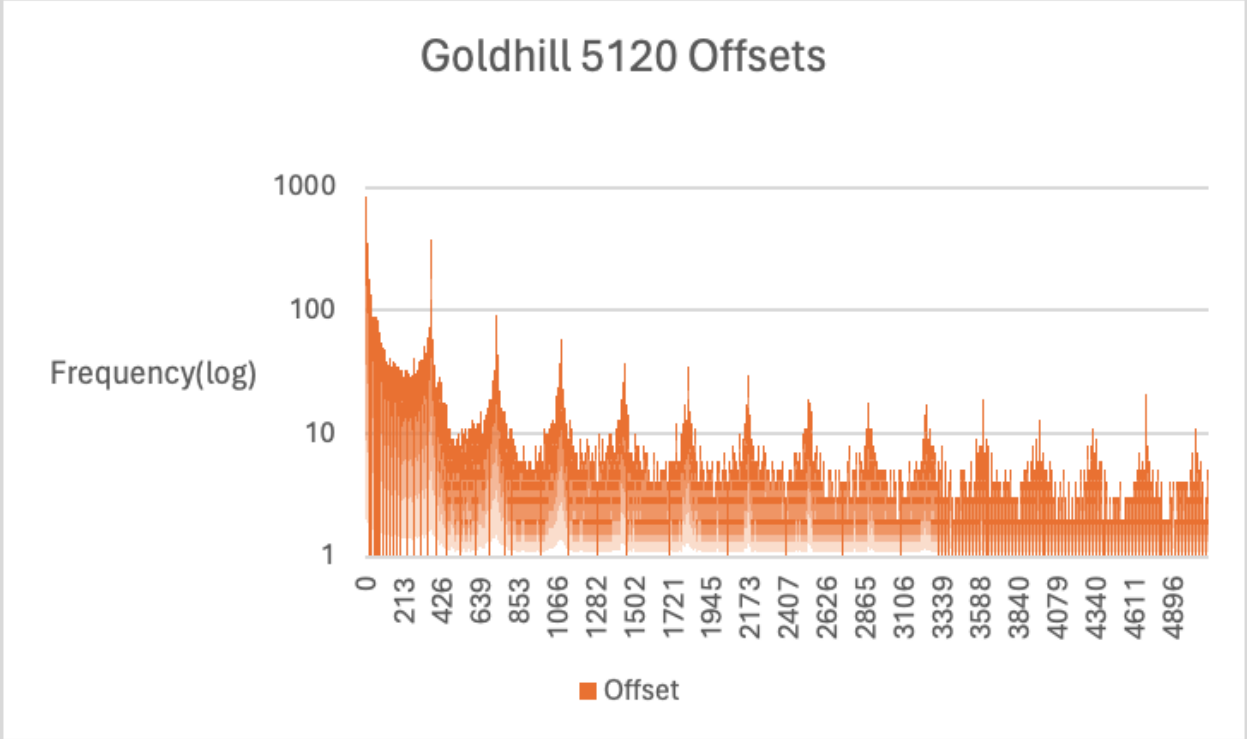
Offset Histograms

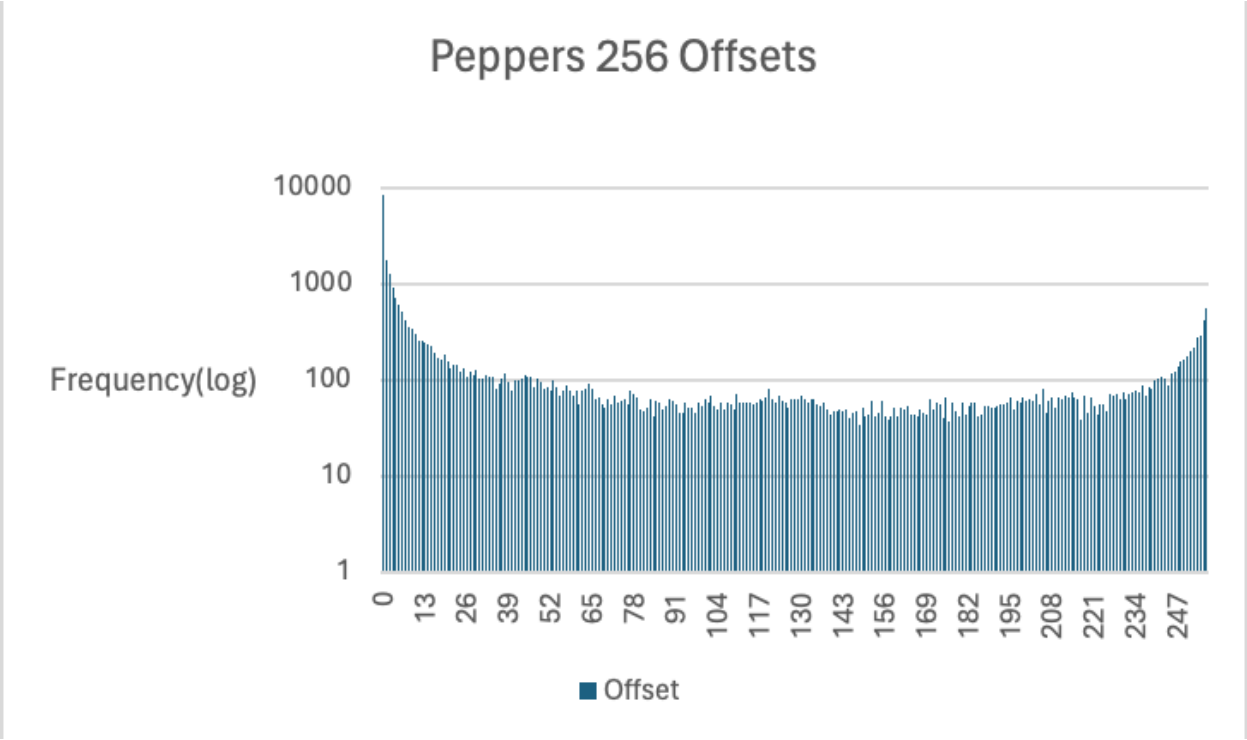


No clear impulses for Goldhill 256, just an initial spike that slowly declines and stabilizes.

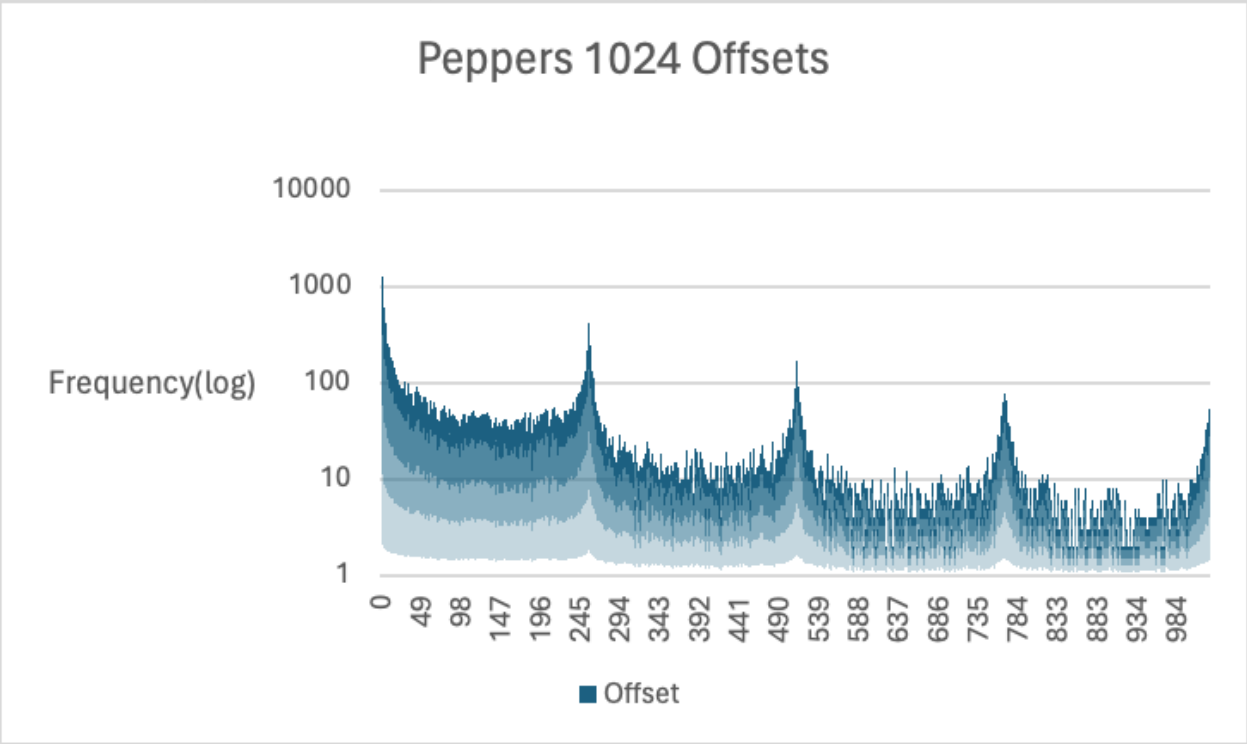


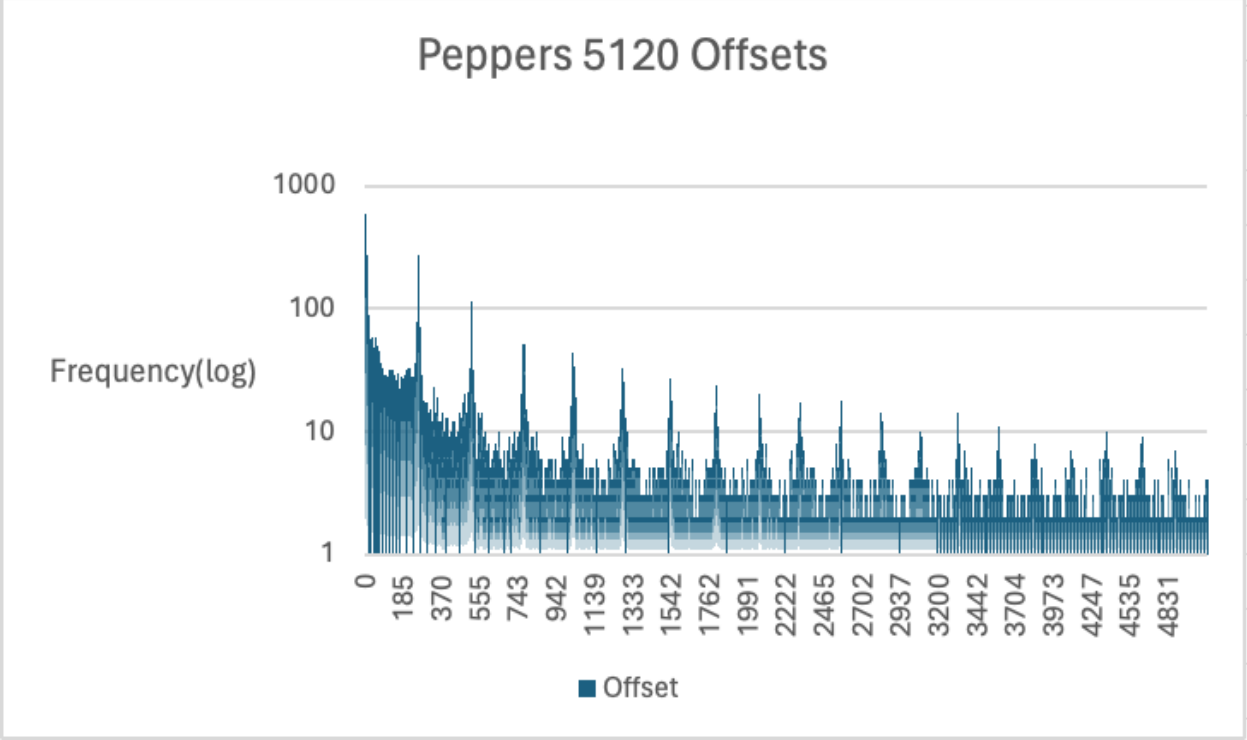


These peaks in Goldhill 1024 and 5120 represent a higher frequency of matches, as the pixels are aligned horizontally across rows.



highest frequencies at the ends, meaning it finds many matching pixels at the edges of the rows.

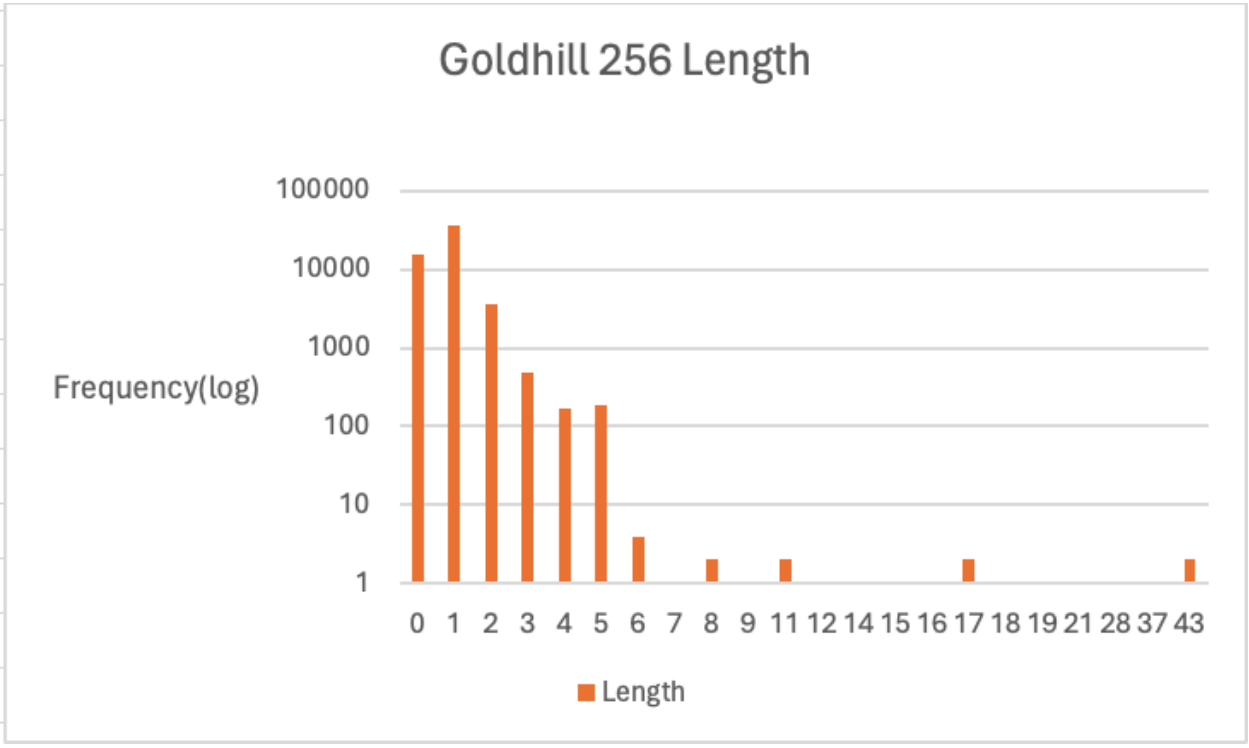


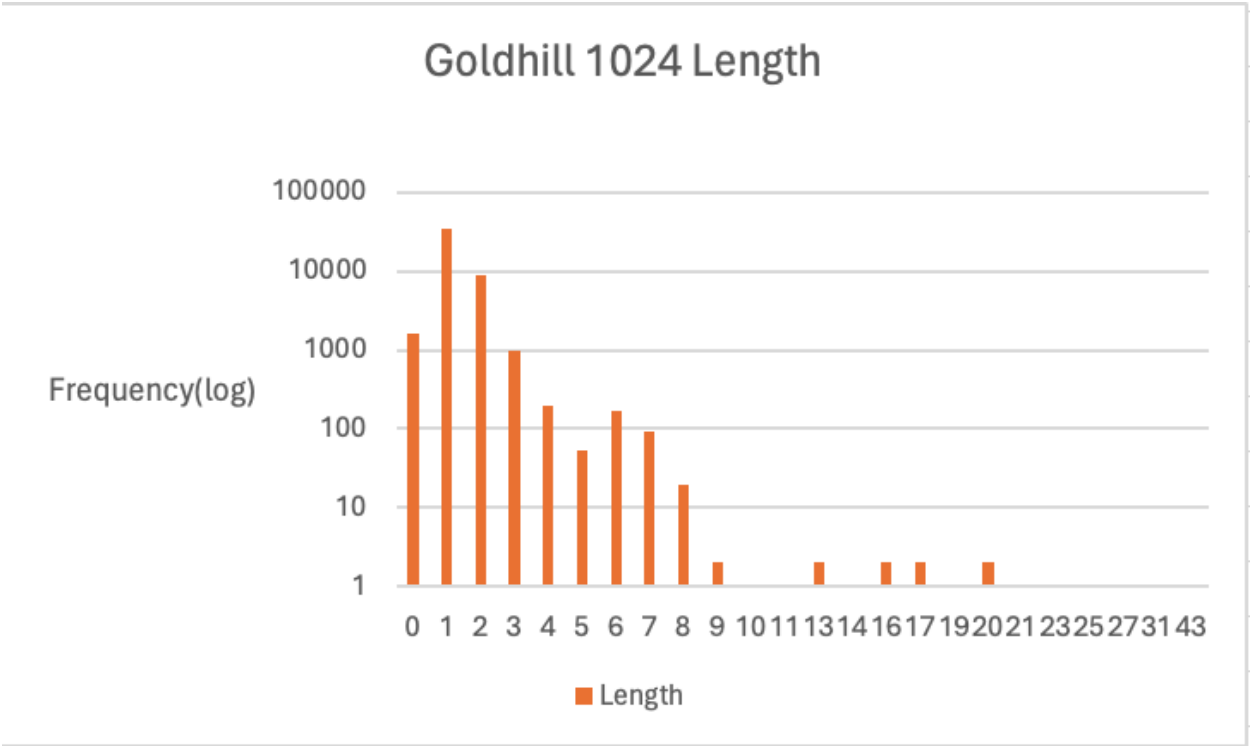


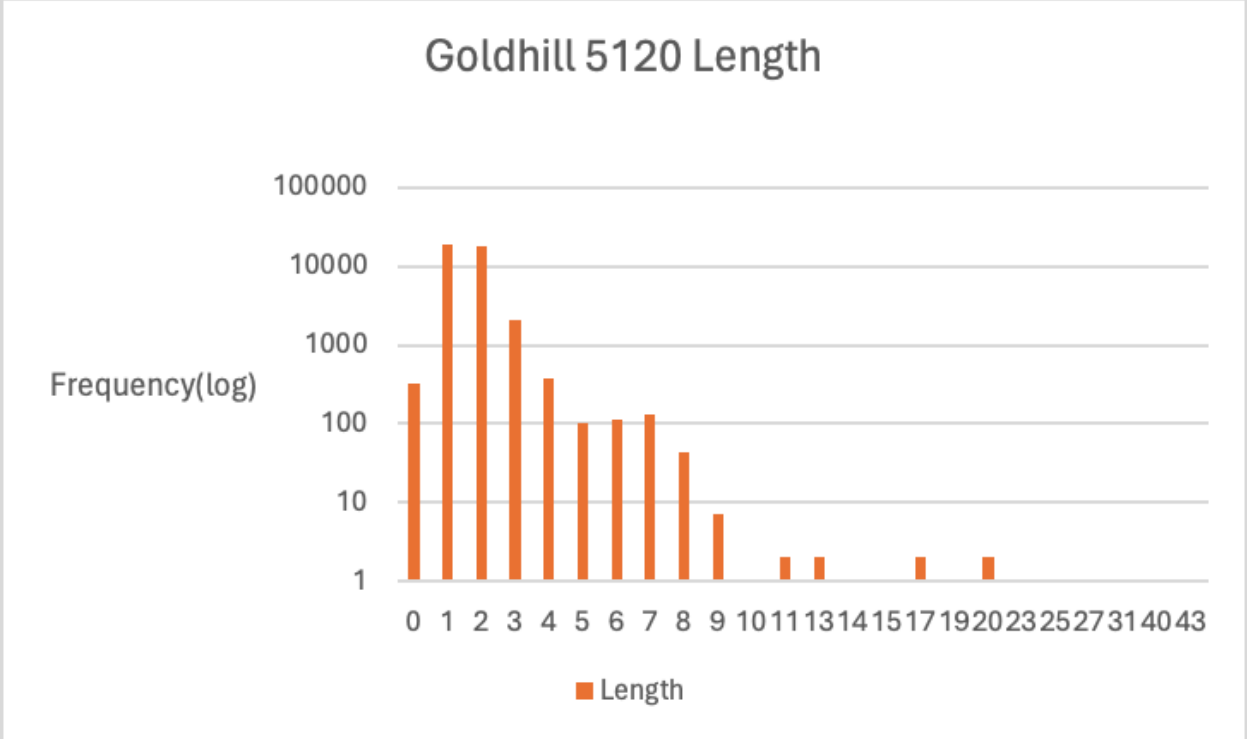
regular impulses in both Peppers 1024 and 5120.

Overall:
The offset distributions of Goldhill and Peppers at 1024 and 5120 show how image width affects the frequency of pixel matches. The width is larger than the offset in the Goldhill 256 offsets, this effect is less noticeable.

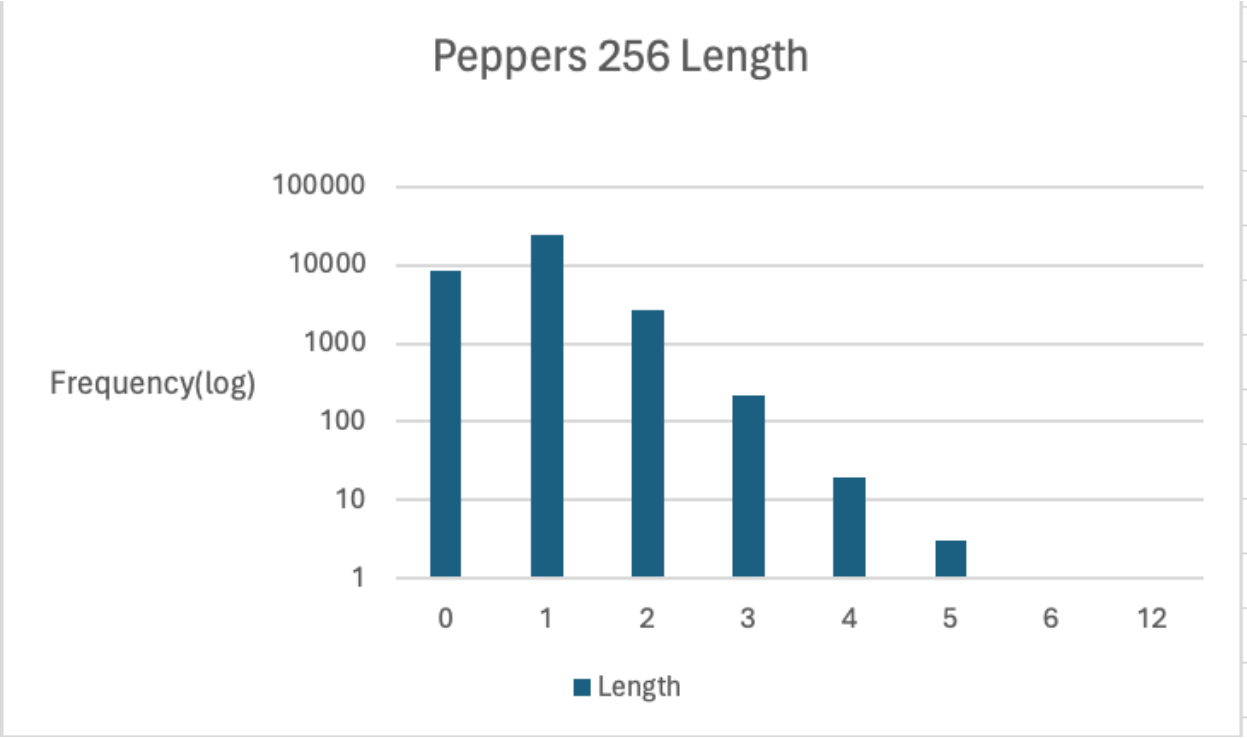
Matching Length Histograms:

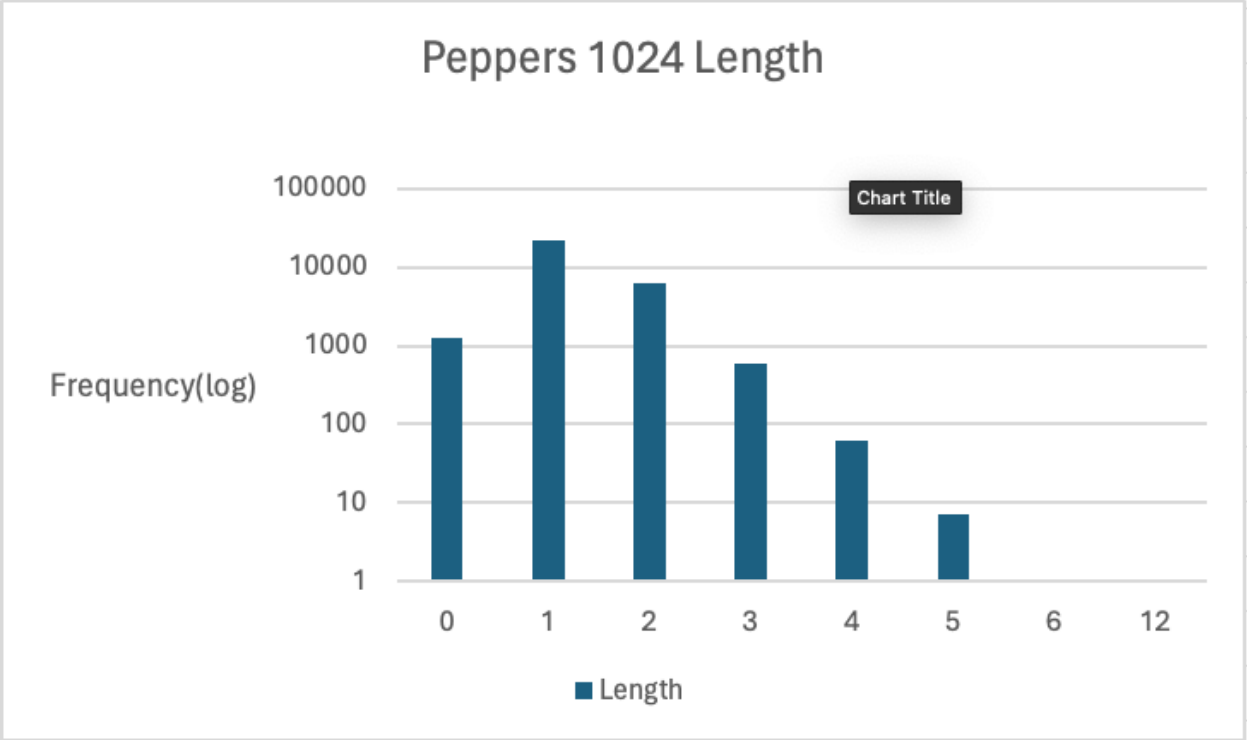


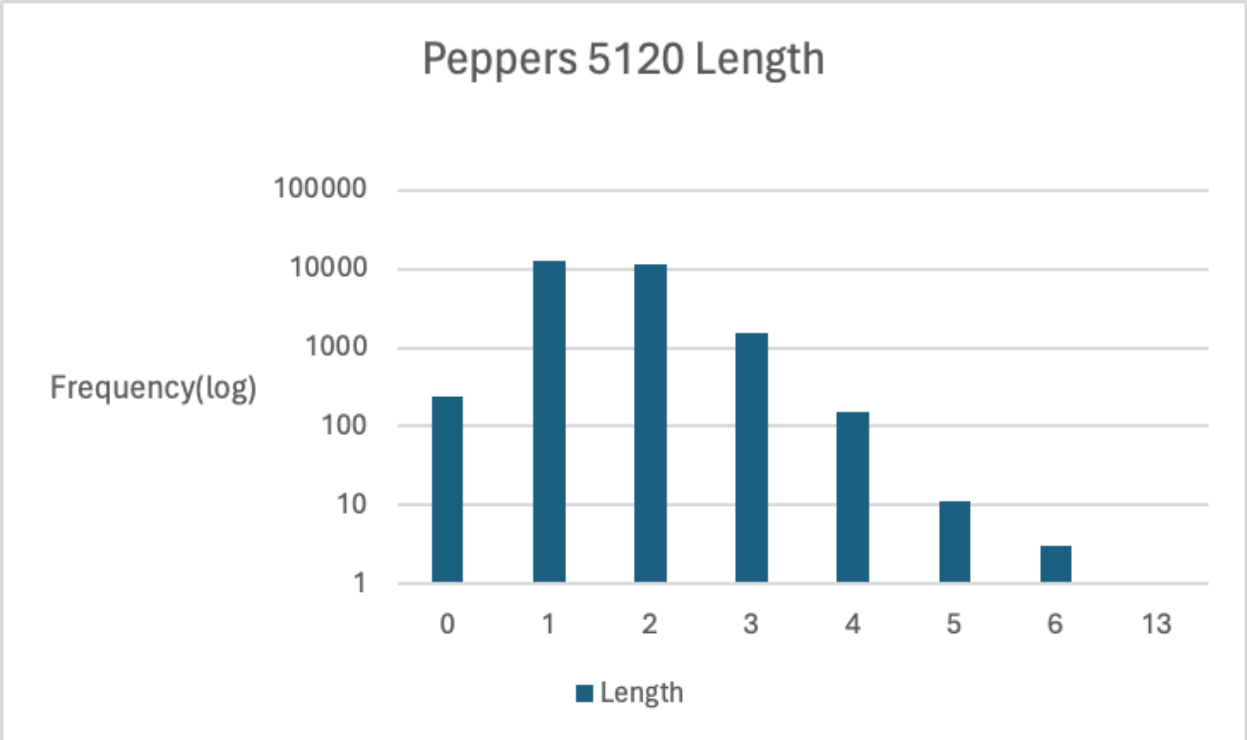




frequency dropping off exponentially as length increases for all Goldhill lengths.







Similar frequency dropping off for Peppers Length

Overall:
both datasets show that the shortest match lengths are the most common, but the Peppers data tends to have more variation and longer matches as the offset size increases. The Goldhill data maintains a steep decline, which implies a lesser occurrence of longer matches.

Average/standard deviation, + encoding/decoding time:

Data	Offset Ave.	Offset Std.	Match Ave.	Match Std.	Encoding Time (s)	Decoding Time (s)
Peppers 256	73.32	90.35	0.85	0.57	0.04	0.01
Peppers 1024	239.44	258.25	1.22	0.56	0.1	0.01
Peppers 5120	1038.01	1301.39	1.57	0.66	0.39	0.01
Goldhill 256	45.36	65.99	0.84	0.76	0.06	0.02
Goldhill 1024	213.22	242.39	1.27	0.81	0.16	0.02
Goldhill 5120	1044.66	1314.11	1.64	0.93	0.61	0.02

Which searching_buffer_size is more suitable?
A searching_buffer_size of 5120is most optimal.