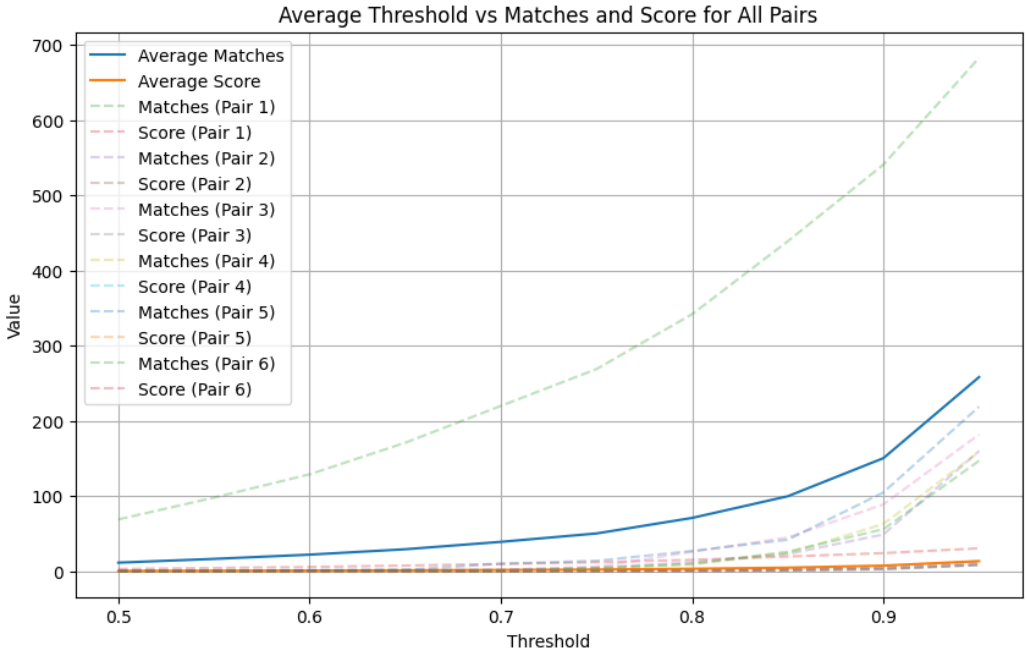
Task 1: Define a reasonable threshold for the similarity between two image-keypoints



Based on these findings from the plot, I'd say that 0.6 seems like the best candidate. While Pair 6 does skew the average, it seems that all the other pairs seem to share a similar trend of increasing more since 0.5 or 0.6. I'd personally go with 0.6 as it would yield more matches, while keeping a balance and not letting it overfit (or, rather, find false matches).

Task 2: Redo the SIFT with all the images (including the 'book1_masked')

The findings: The image with book1 is kind of similar to the book2 image, but it clearly shows that the masked book is the same image (or is supposed to be) as book1. 35% was the finding. I'd say that by omitting the other results that are basically rounding errors, we can conclude that book1 is most likely the same as book1_masked. I, unfortunately, do not have the space to fit the pictures all into this small PDF file, so I will only leave the best result which is book1 + book1_masked. All code and results can be found in the jupyter notebook.