Comparing the similarity of pictures based on varying the “NumActiveColumnsPerInhArea” variable, while using localInhibition

To use the NumActiveColumnsPerInhArea parameter, the value of localAreaDensity has to be set to a negative value

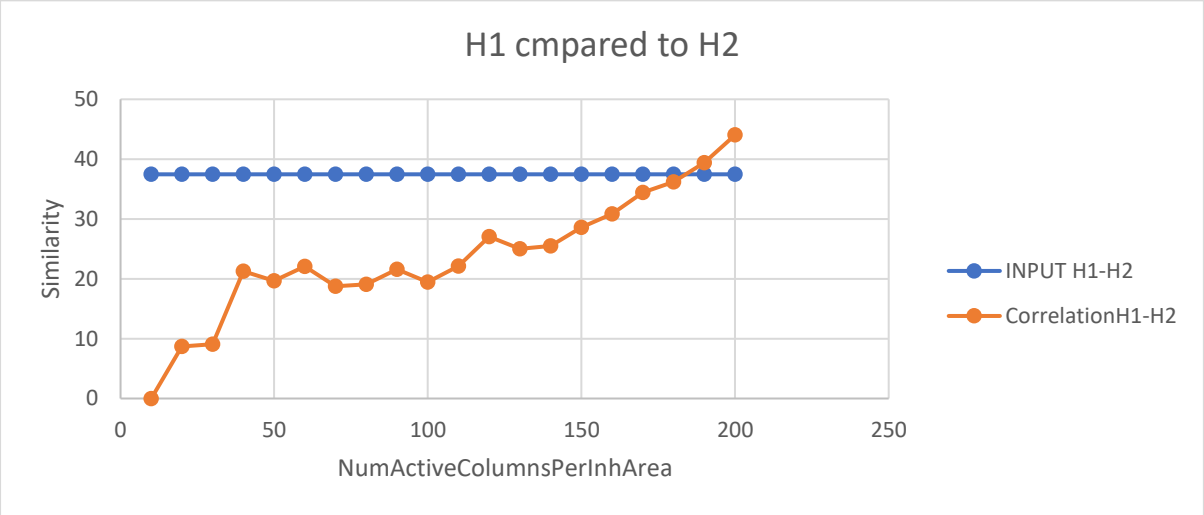

Figure 1 : Comparing Hexagon1 to Hexagon2, varying the NumActiveColumnsPerInhArea from 10 to 200

H1 H2

The similarity is increased as the NumActiveColumnsPerInhArea is increasing. Increment of similarity of H1_H2 is shows that the HTM sees the images of a same category more and more similar, as the NumActiveColumnsPerInhArea increases

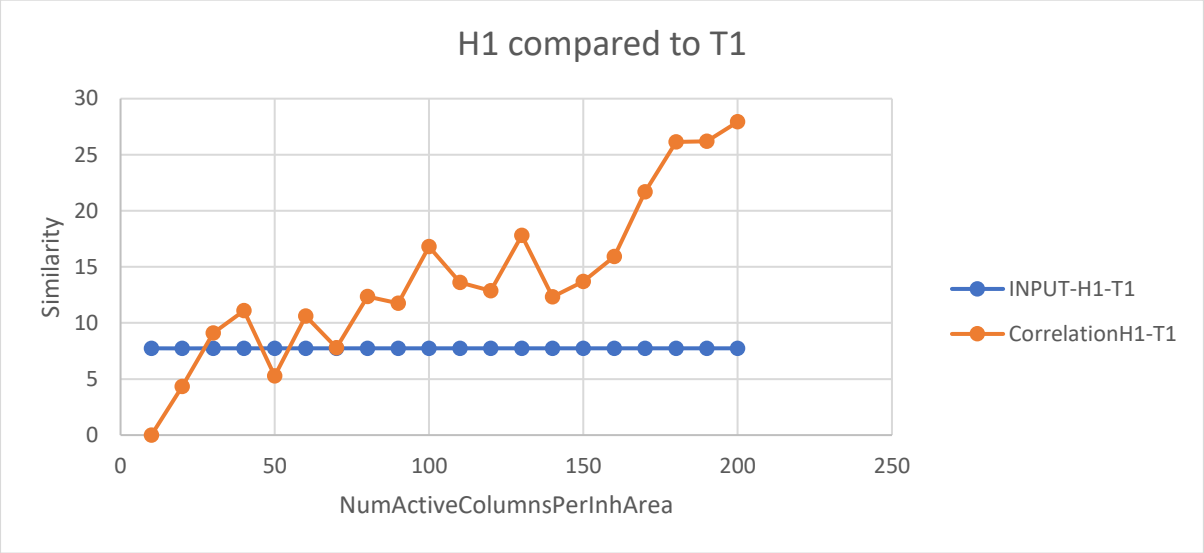

Figure 2 : Comparing Hexagon1 to Triangle2, varying the NumActiveColumnsPerInhArea from 10 to 200

H1 T1

When we compare Hexagon1 to Triangle1 as the NumActiveColumnsPerInhArea is increased the similarity is also increased. This is not a good news and means HTM considers pictures of different categories more and more similar to each other as NumActiveColumnsPerInhArea is increased, and this is not desired. Specially between NumActiveColumnsPerInhArear of 150 and 190 that the similarity increases with a higher rate.

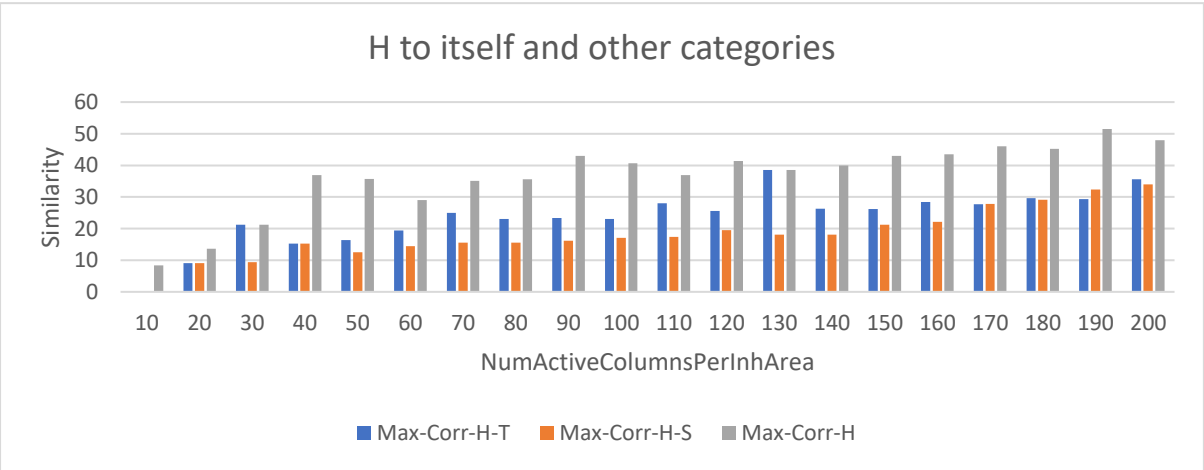

Figure 3: the Max-similarity of Hexagon category to itself compared to Max-similarity of Hexagon to other categories varying NumActiveColumnsPerInhArea from 10 to 200

comparing the max-micro-similarity of Hexagon category to max-macro-similarity between Hexagon category and other categories, could be a good approach to choose the best NumActiveColumnsPerInhArea parameter among 10-200. Because here we can have a clearer sight on how HTM system is functioning. For example, for NumActiveColumnsPerInhArea=10 the similarity is not that high, but the safe margin (the difference of Max-micro-Corr-H compared to Max-macro-Corr-H to other categories) is 0. But for example, as we consider NumActiveColumnsPerInhArea=30, the Max-Correlation of H to itself is increased to 22% where Max Correlation of Hexagon to Triangle is also 22%, this means that due to a safe margin of 0, the system in this point is not reliable at all.

So the reliability factors could be a high similarity of a category members to themselves, and lowest possible similarity to other categoriy members. As it can be seen in NumActiveColumnsPerInhArea=90,150,170 where Max-micro-Corr-H is interestingly high and Max-macro-Corr-H, is sufficiently low, the most sufficient amount of NumActiveColumnsPerInhArea can be observed.

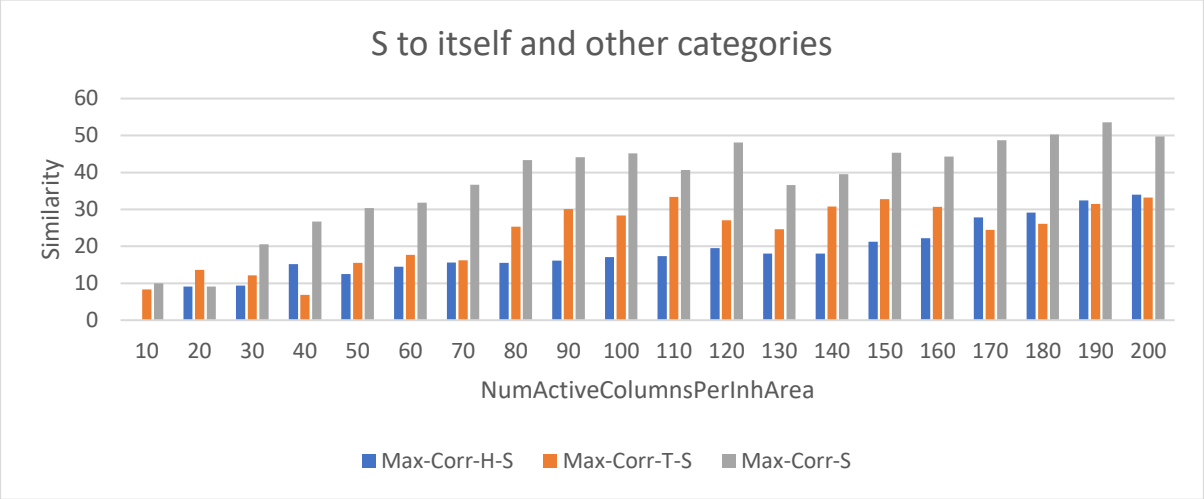

Figure 4: the Max-similarity of Straight cross category to itself compared to Max-similarity of Straight cross to other categories varying NumActiveColumnsPerInhArea from 10 to 200

similarity of Straight cross category to itself and other categories, shows us the increase of similarity to itself and other categories, as the NumActiveColumnsPerInhArea is increased, our desired NumActiveColumnsPerInhArea is actually a value, where the similarity of the Straight cross category to itself is high, and at the same time the similarity to other categories are as low as possible. Here the NumActiveColumnsPerInhArea= 120,170 or 180 are the best possibilities for us.

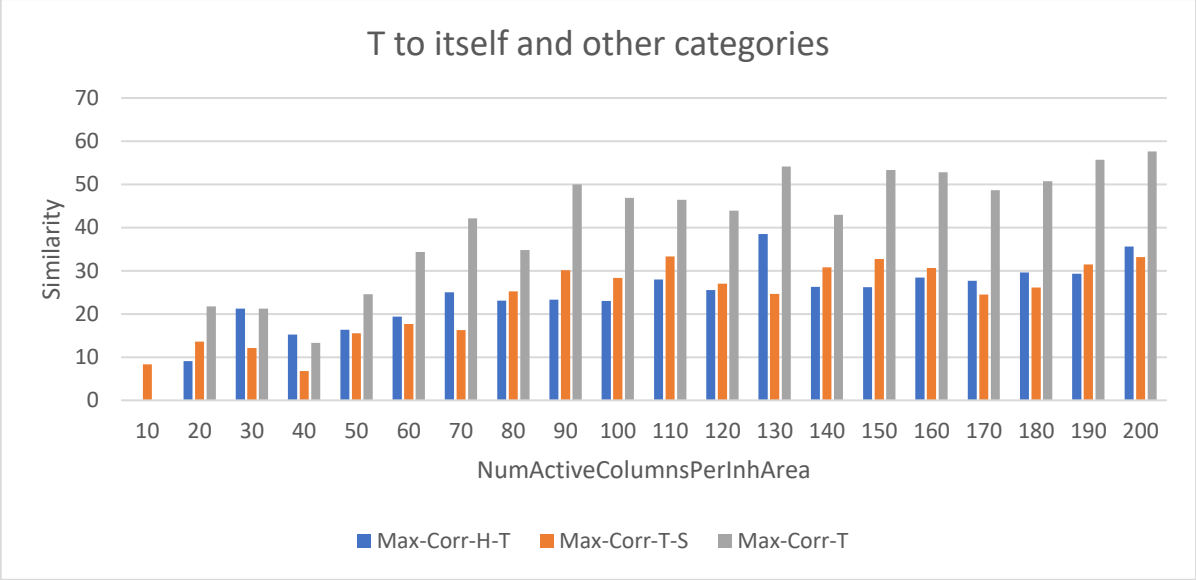

Figure 5: the Max-similarity of Triangle category to itself compared to Max-similarity of Triangle to other categories varying NumActiveColumnsPerInhArea from 10 to 200

As it can be seen raising NumActiveColumnsPerInhArea up to 50 is not providing us an acceptable result, as in NumActiveColumnsPerInhArea=20 the safe margine(the difference of Max-micro-Corr-T compared to Max-macro-Corr-T to other categories) is not large enough. And in NumActiveColumnsPerInhArea=30 or 40 , we are experiencing overlaps. But in NumActiveColumnsPerInhArea=70,160 or 170, the correlation of Triangle to itself is high enough, and also a reliable safe margin is provided.

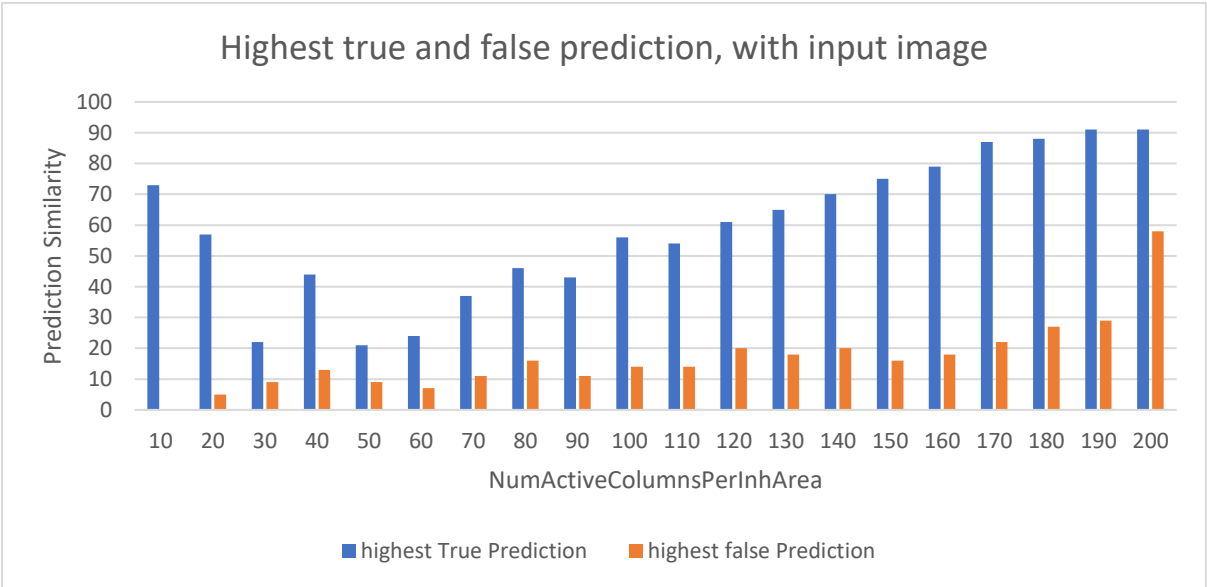

Figure 6: comparing the true prediction percent to the highest false prediction percent, in order to have a sight on reliability of prediction part of the system varying NumActiveColumnsPerInhArea from 10 to 100

The best prediction of 100% with a 0 false prediction happens in NumActiveColumnsPerInhArea=10 , but we can consider this extremely correct and accurate prediction, due to not properly educated system(probably when the NumActiveColumnsPerInhArea=10 there is not enough amount of active cells in SDR) so we don't consider it as a good possibility. On the other hand, the prediction part of the system is showing great results in NumActiveColumnsPerInhArea=150,160,170 and 180. Also the

safe margin is really good.

conclusion

when NumActiveColumnsPerInhArea =170 we can not only see a good performance of system in macro and micro correlations, but also a well doing reliable prediction. So we can consider the NumActiveColumnsPerInhArea= 170 as the best possible NumActiveColumnsPerInhArea value for a range between 10_200 when the localInhibition and NumActiveColumnsPerInhArea parameters are used.

Comparing the similarity of pictures based on varying the “local area density” variable, while using GlobalInhibition

In order to get advantage of globalInhibition, we have to set its valu to True, and because of using local areaDensity, we should turn NumActiveColumnsPerInhArea to a negative value

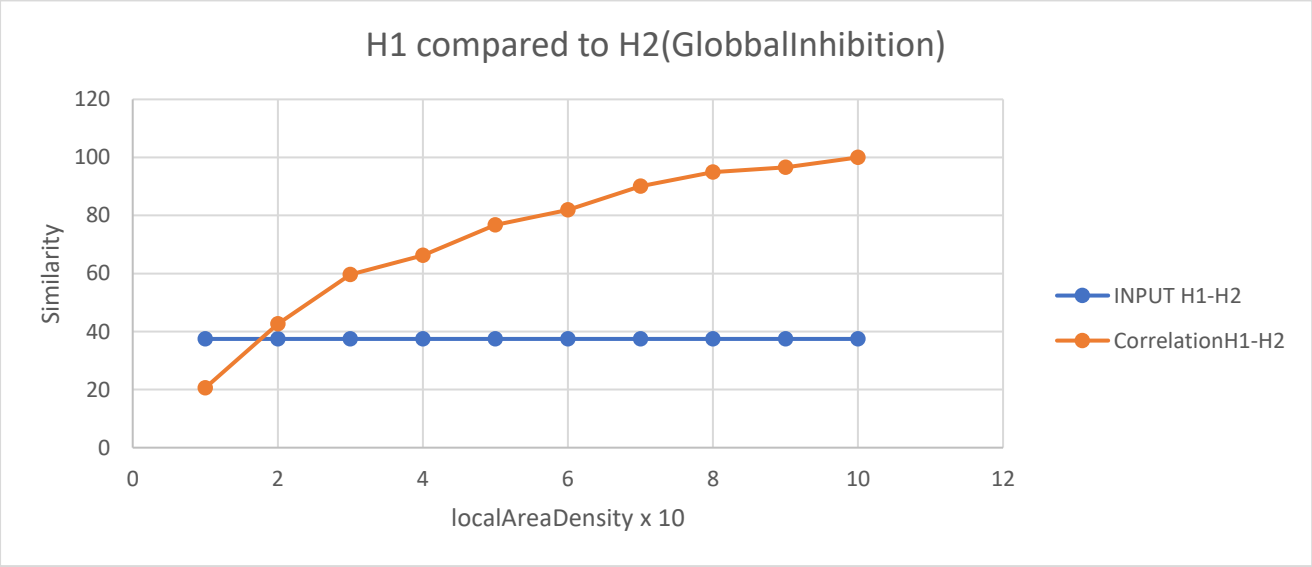

Figure 7 : Comparing Hexagon1 to Hexagon2, varying the localAreaDensity from 10 to 100

By increasing the localAeaDensity the H1 and H2 are looking more and more similar for the system, but even in the beginning of this increment the speed of similarity growth of these two is really high.

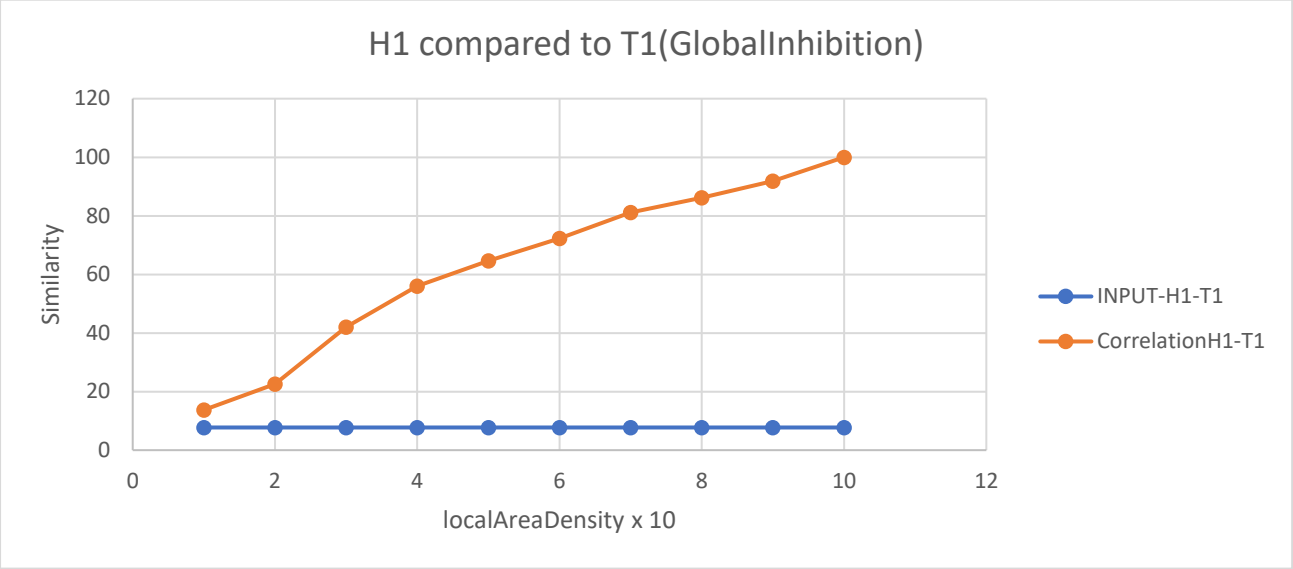

Figure 8 : Comparing Hexagon1 to Triangle 1, varying the localAreaDensity from 10 to 100

By increasing the localAeaDensity the H1 and T1 are looking more and more similar for the system, though the similarity growth I the beginning is not as high as the similarity growth in last part (comparing H1-H2) this means the higher localAeaDensity values couldn't be a good choice , as the system sees different shapes similar to eachother

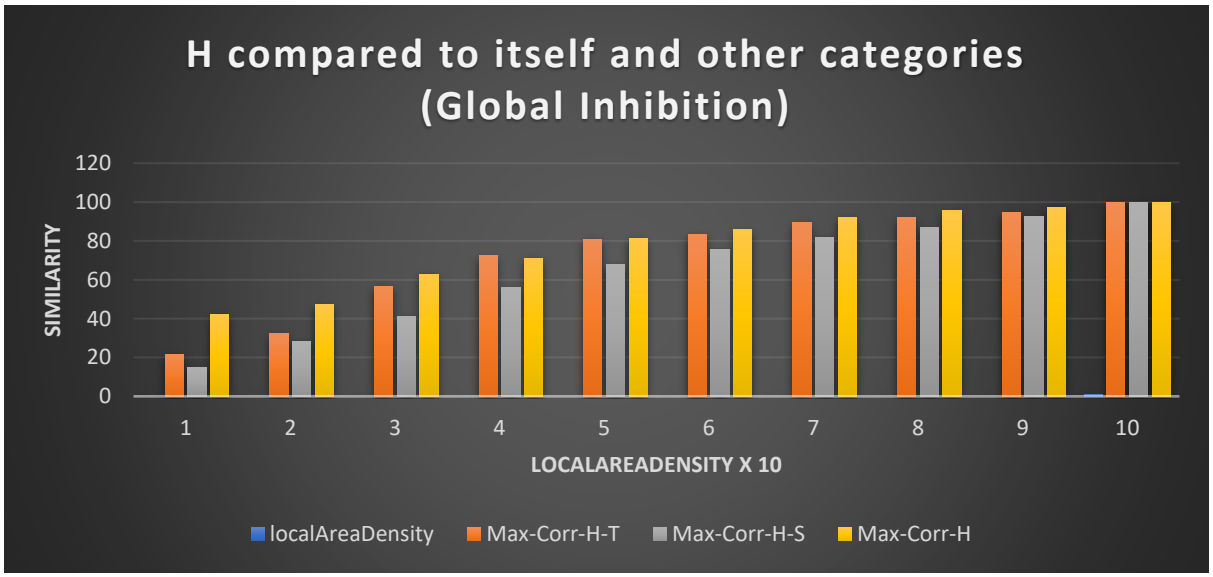

Figure 9: the Max-similarity of Hexagon category to itself compared to Max-similarity of Hexagon to other categories varying localAreaDensity from 0.1 to 1

Though the increment of localAeaDensity leads to higher similarities of hexagon to itself, its also increasing similarity of other categories to hexagon too. This means a lower reliability in higher localAeaDensity values. In this case the best amounts of localAeaDensity are 0.1 and 0.2 where an almost good safe margin is observable

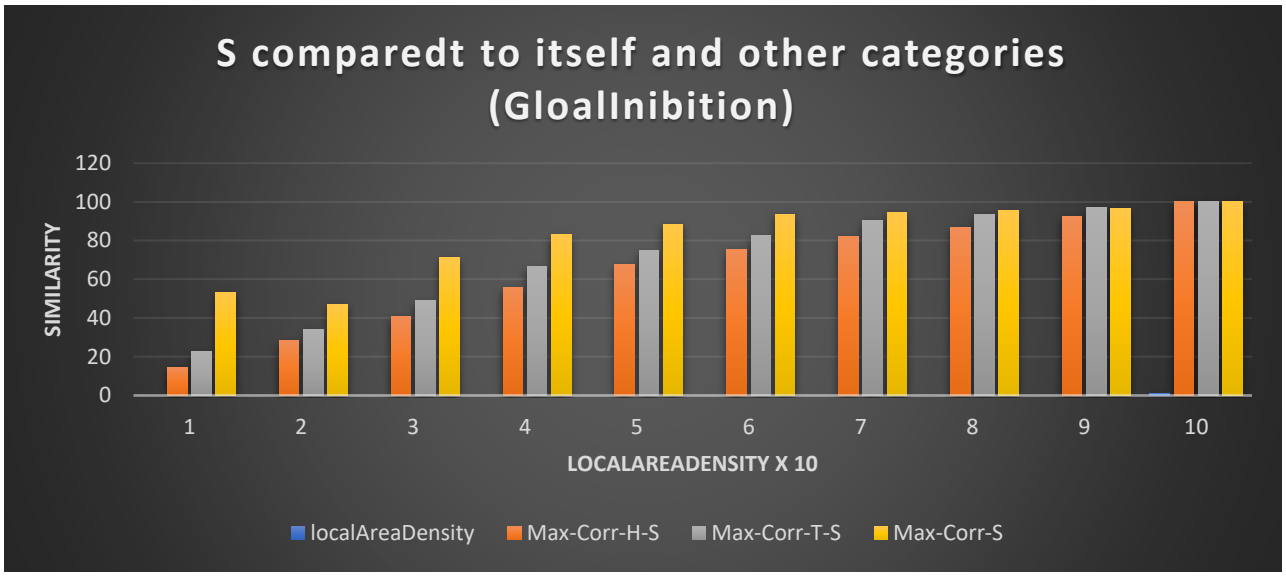

Figure 10: the Max-similarity of Straigh cross category to itself compared to Max-similarity of Straight cross to other categories varying localAreaDensity from 0.1 to 1

The same unwanted growth of similarity of other categories to Straight cross category is observable here too. The best safemargines (the difference of Max-Corr-S compared to other categories) happens when the localAeaDensity is equal to 0.1 and 0.2 and 0.3

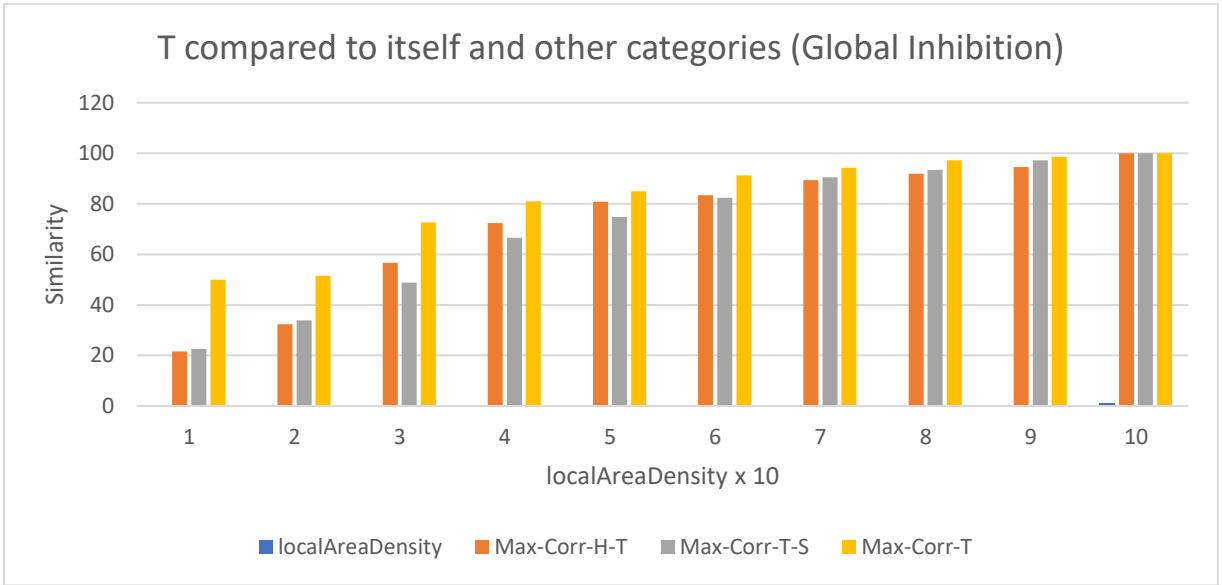

Figure 11: the Max-similarity of Triangle category to itself compared to Max-similarity of Triannge to other categories varying localAreaDensity from 0.1 to 1

The best safe margins are happening in localAeaDensity of 0.1 and 0.2 and 0.3 here too.

Actually the higher the localAreadensity, the higher amount of sparsity in our SDRs , this means we get more 1s. the localAeaDensity of over 60 reduces the ability of system to produce an extremely unique SDR for each image, this couses the system to see different shapes pretty similar to eachother, as the produced SDR of each image is similar to other ones

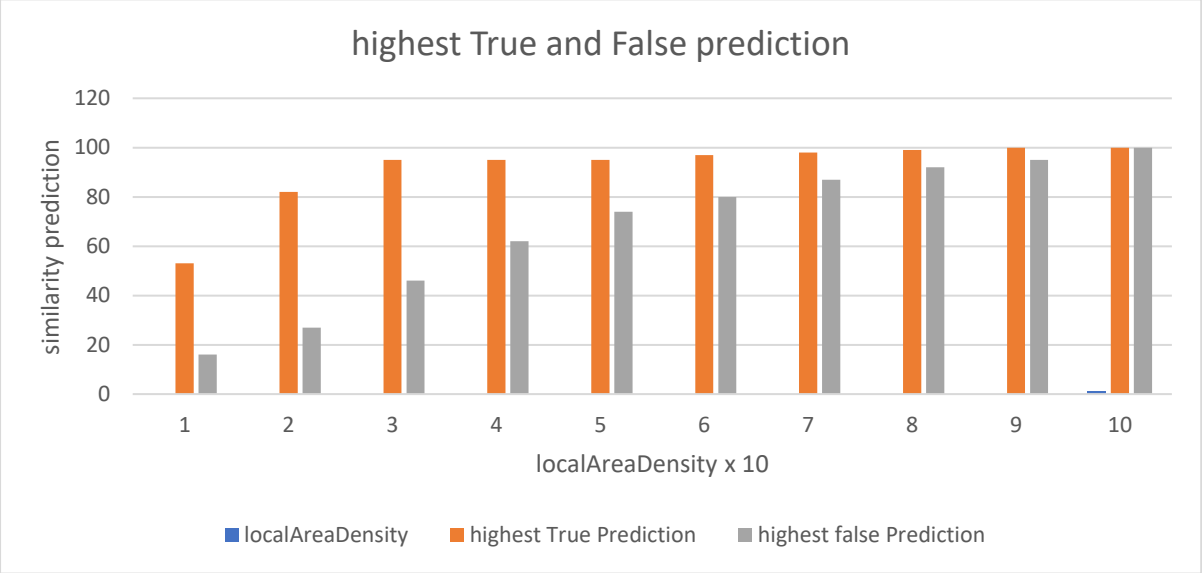

Figure 13: comparing the true prediction percent to the highest false prediction percent, in order to have a sight on reliability of prediction part of the system varying localAreaDensity from 0.1 to 1

The higher our localAeaDensity, the higher the percentage of false prediction.

this figure shows us too that we should avoid higher localAeaDensity values. And we gotta concentrate on higher margins as higher reliability is more desired
so the best local area denity for a good prediction would also be : 0.1 – 0.2 – 0.3