# Experimental results

The algorithm includes various parameters that can be tuned:

– Affects the rigidity of the displacement field. The higher is, the displacement field is more restrained.

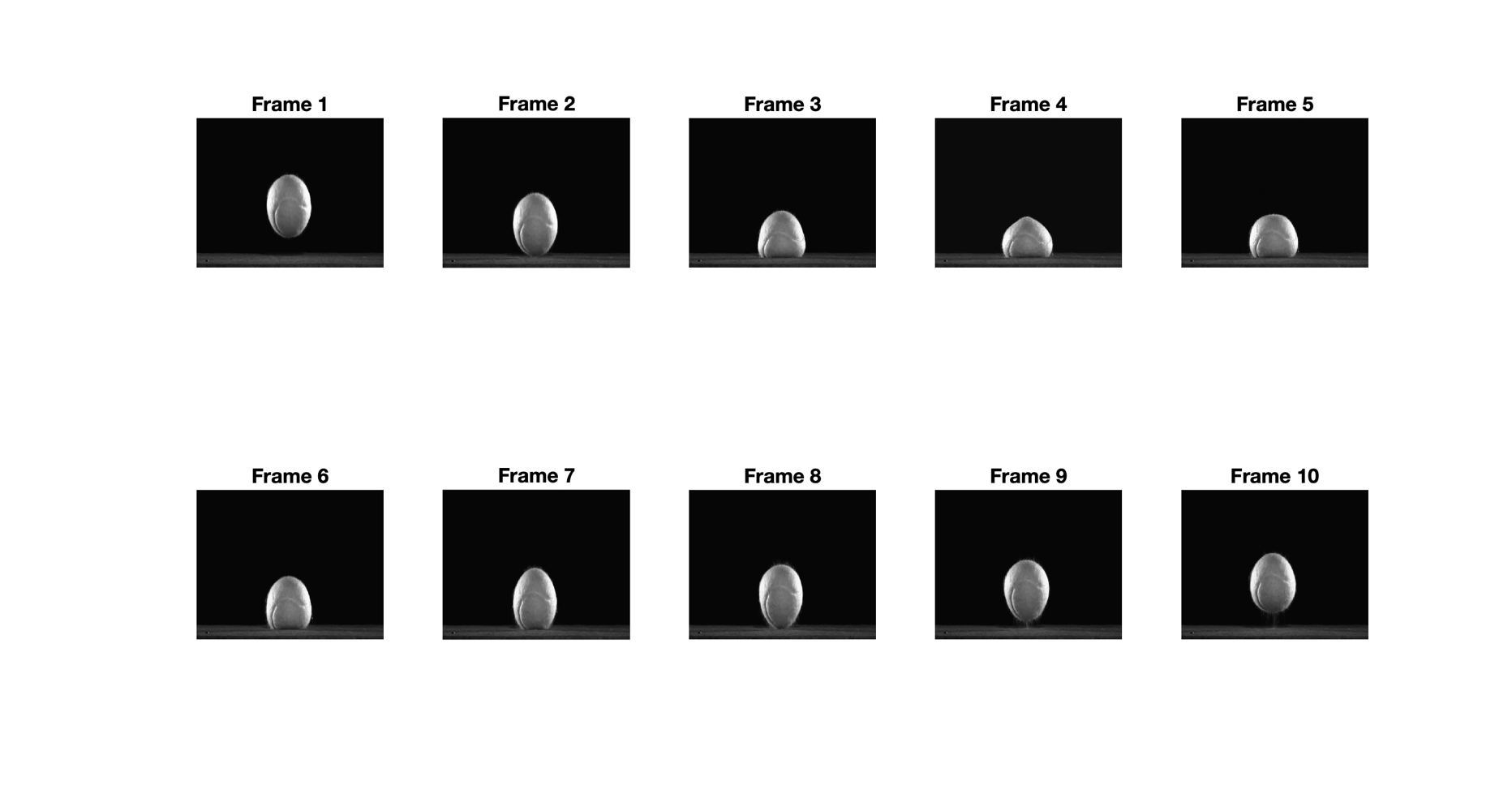
**Taylor Series Order** – The data term is linearized with a Taylor series to make the model convex. The higher the order the, the more accurate the series is to the original data term.

**Scale Levels** – The algorithm works on different scale levels to accelerate convergence and utilizing course information correlation (focusing on general high frequency details such as shapes as opposed to fine details such as texture).

**TV Order** – The power order over the absolute of image derivatives. First order is ordinary TV. Higher order derivatives allow for smooth-inducing and discontinuities regularizations as well as being less dependent on initial alignment between images.

The following graphs explore the impact of each parameter over the resulting IOU of a short 10-frame video of a tennis ball falling at high speed.

Below are the frames tested on. The challenge posed by the video is misalignment of the target between consecutive frames as well as deformities in its shape.



**Effect of**

We’ve tested different values for = {1, 10, 100, 1000, 10000}

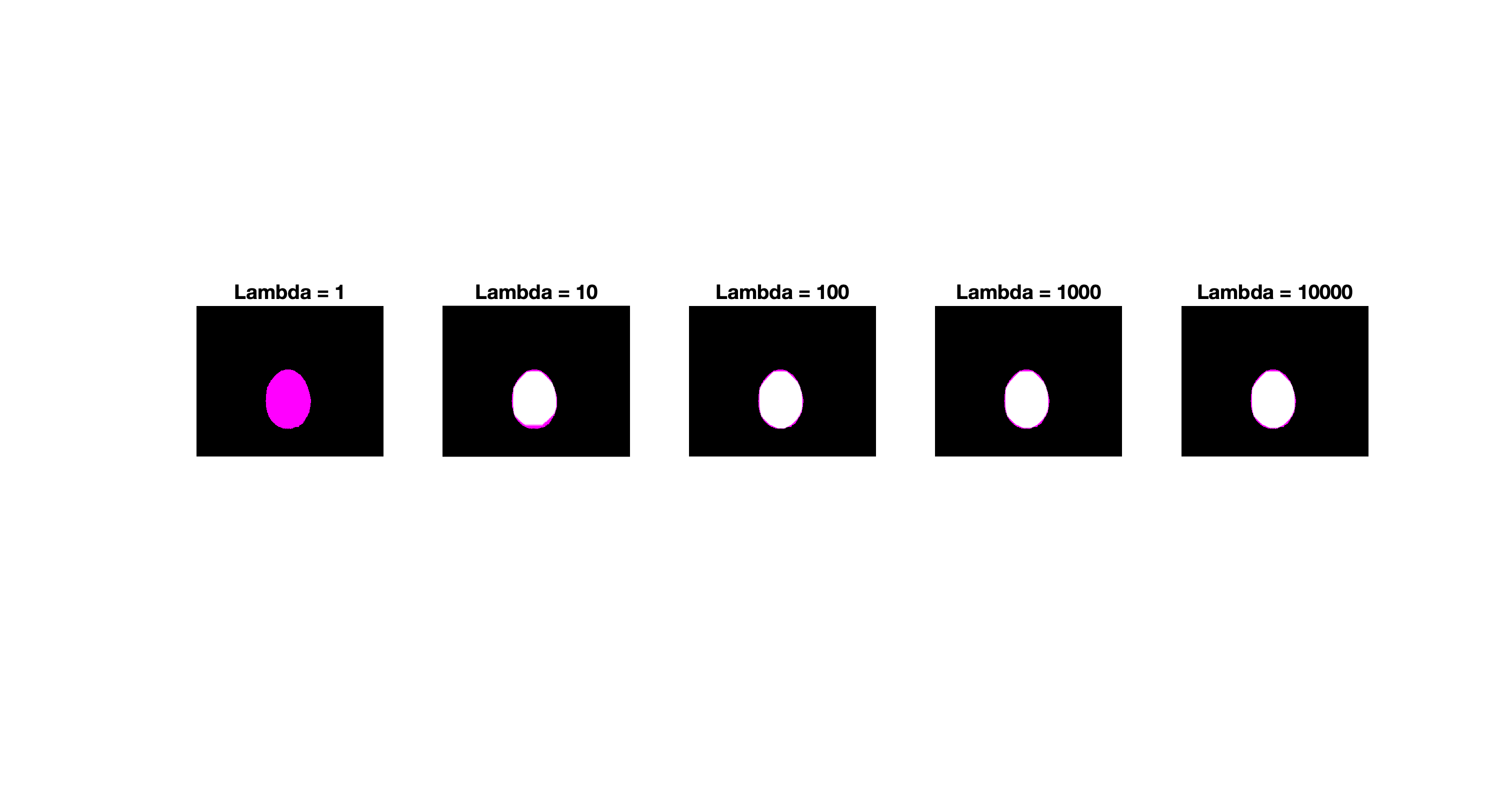
Other fixed values:

Taylor Series Order = 10

Scale Levels = [16, 8, 4, 2, 1]

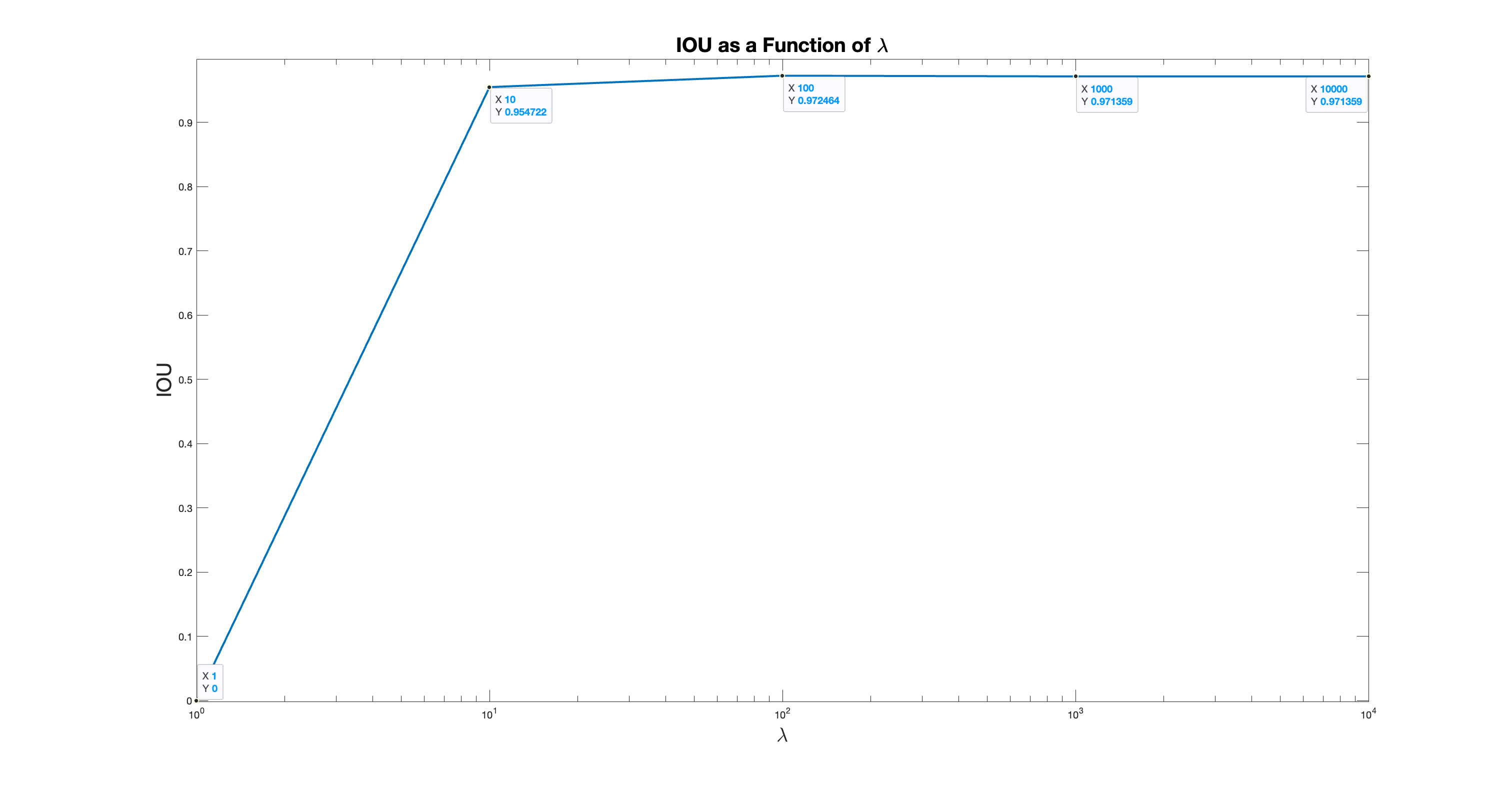
TV Order = 2nd order

The following is the IOU between the algorithm’s mask and ground truth mask in the last image for different :



(White is overlap. Magenta is GT mask. Green is algorithm mask)

The following is a graph of IOU value as a function of :



After we reach a plateau, however the algorithm takes longer to converge the higher is. Therefore, for later tests the value is taken.

**Effect of Taylor Series Order**

We’ve tested different Taylor Series Order values = {1, 3, 5, 7, 10, 15}

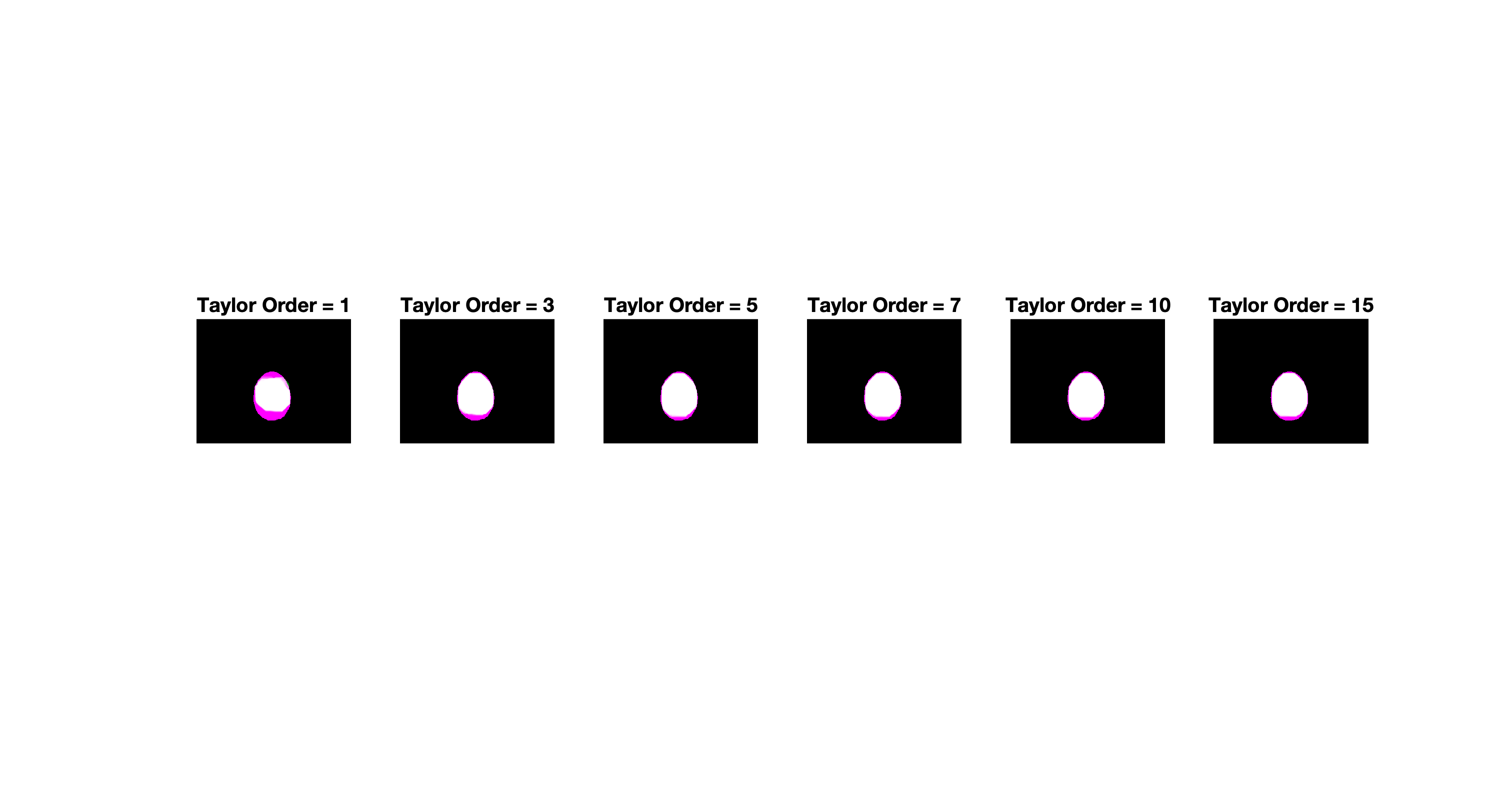
Other fixed values:

= 10

Scale Levels = [16, 8, 4, 2, 1]

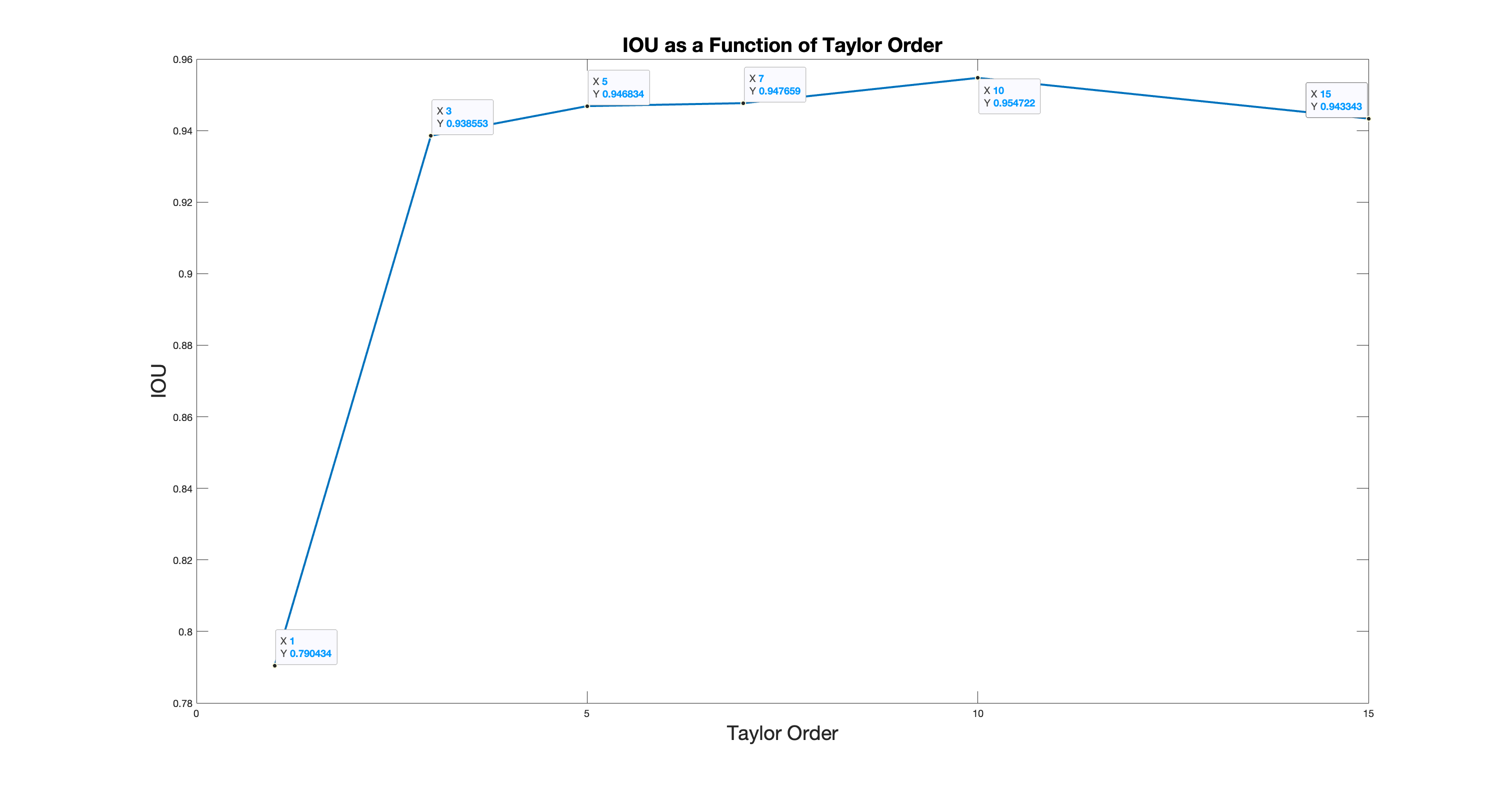
TV Order = 2nd order

The following is the IOU between the algorithm’s mask and ground truth mask in the last image for different Taylor Series Orders:



In the algorithm, when convergence of the data term is checked, the Taylor approximation is used. The higher the order the more precise the convergence is. We can see that the resulting masks are more accurate the higher the order, but with diminishing returns as opposed to the time of convergence.

The following is a graph of IOU value as a function of the Taylor Series order:



Best results were received for Taylor Series Order = 10, and that is the value we fixed for the rest of the tests.

**Effect ofScale Levels**

We’ve tested different pyramid maximum scale levels= {1, 2, 4, 8, 16}

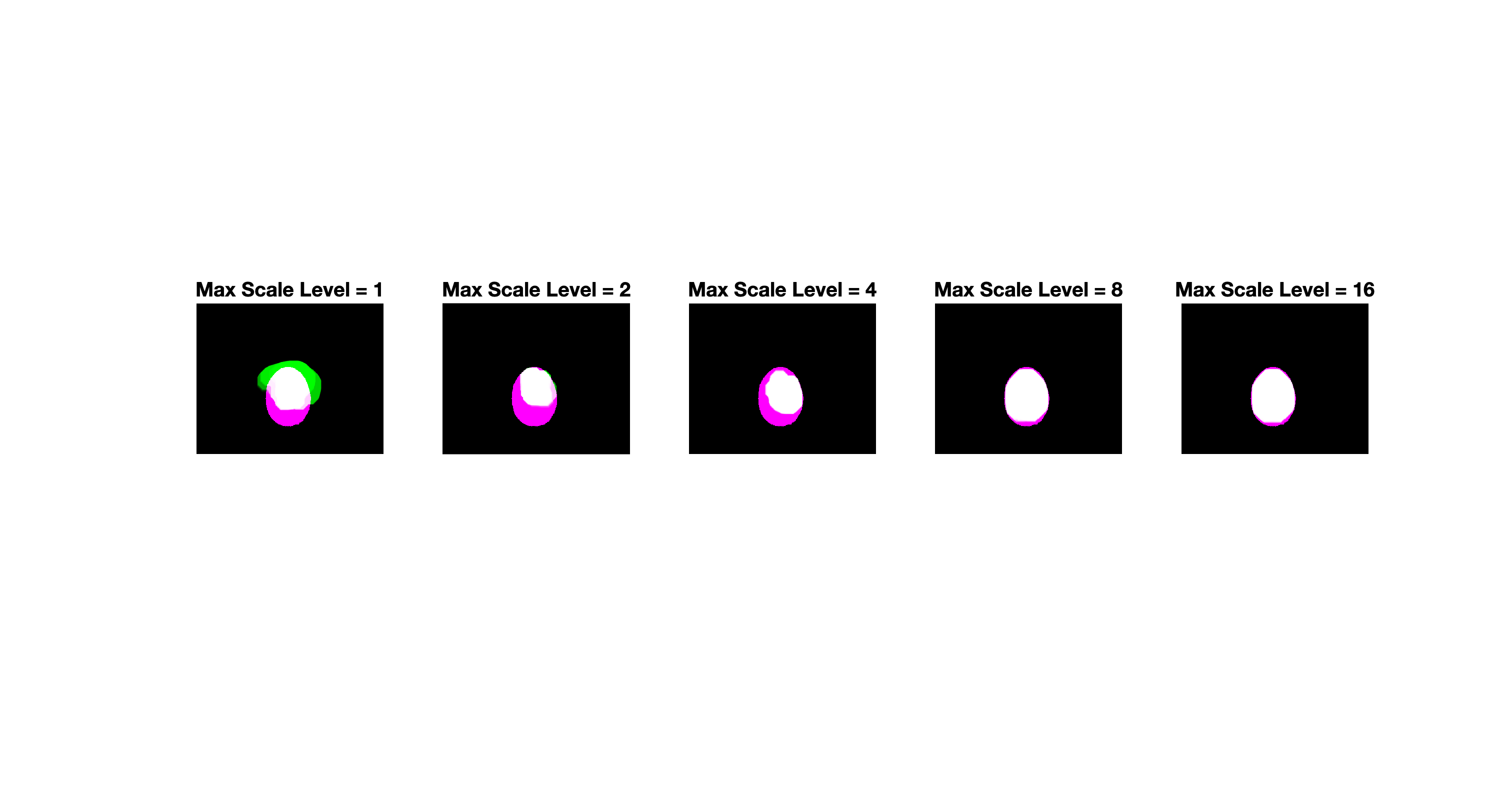
Other fixed values:

= 10

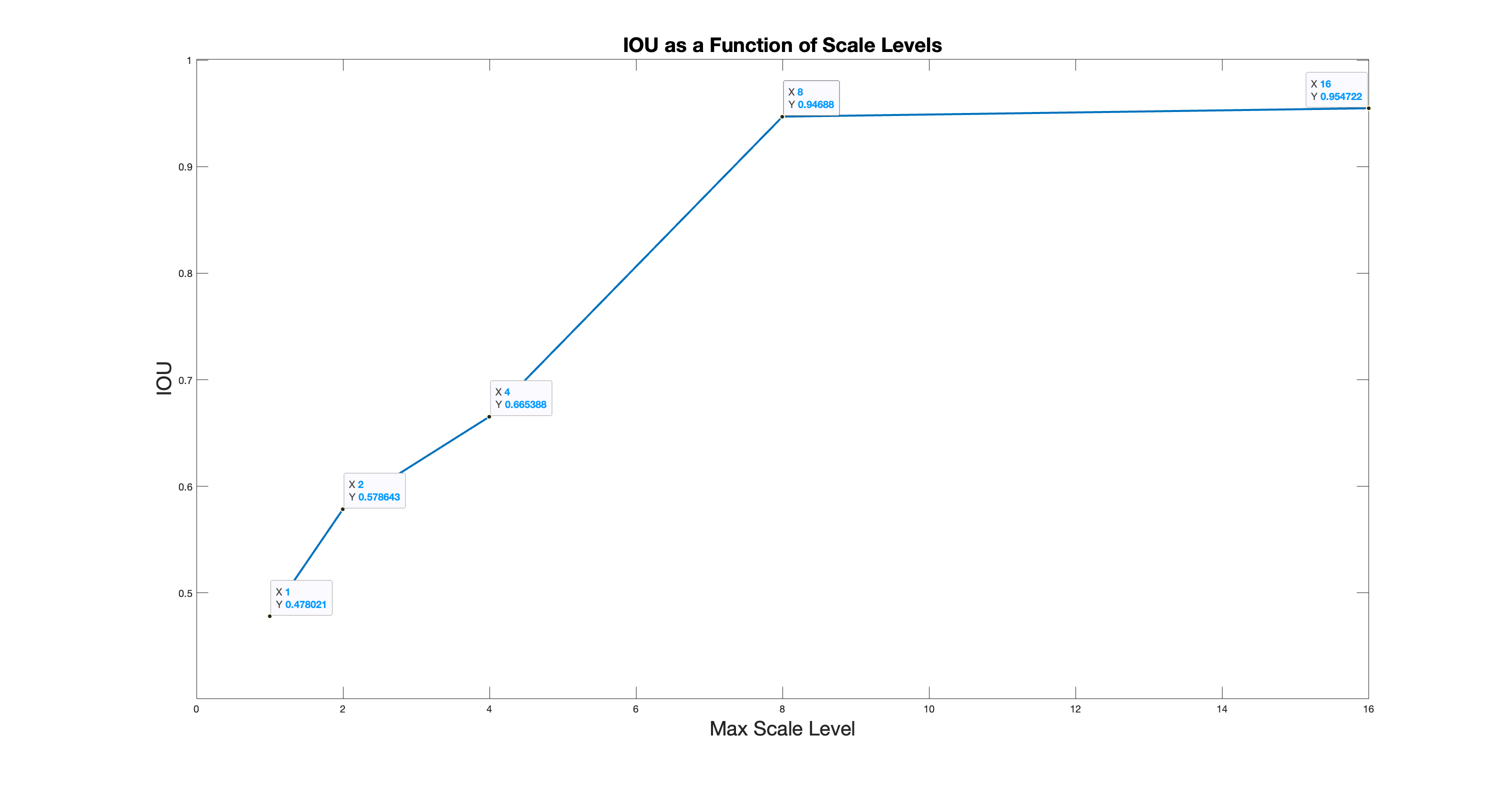
Taylor Series Order = 10

TV Order = 2nd order

The following is the IOU between the algorithm’s mask and ground truth mask in the last image for different sets of scale levels (each one starts from the highest one and dowsamples by a factor of 2 until reaching 1):



And the IOU values themselves as a function of maximum scale level:



Down-sampling the images with higher scales levels helps the iterative process by first matching the images using their “low frequency” features, such as shape and general pixel values.

Once this is done, the images are better aligned, which in turn leaves the rest of the process focus on the finer details such as texture.

As mentioned before, initial alignment is important to the success of TV flow methods.

**Effect ofTV Order**

We’ve tested different TV orders= {1st Order, 2nd Order, 3rd Order, 4th Order}

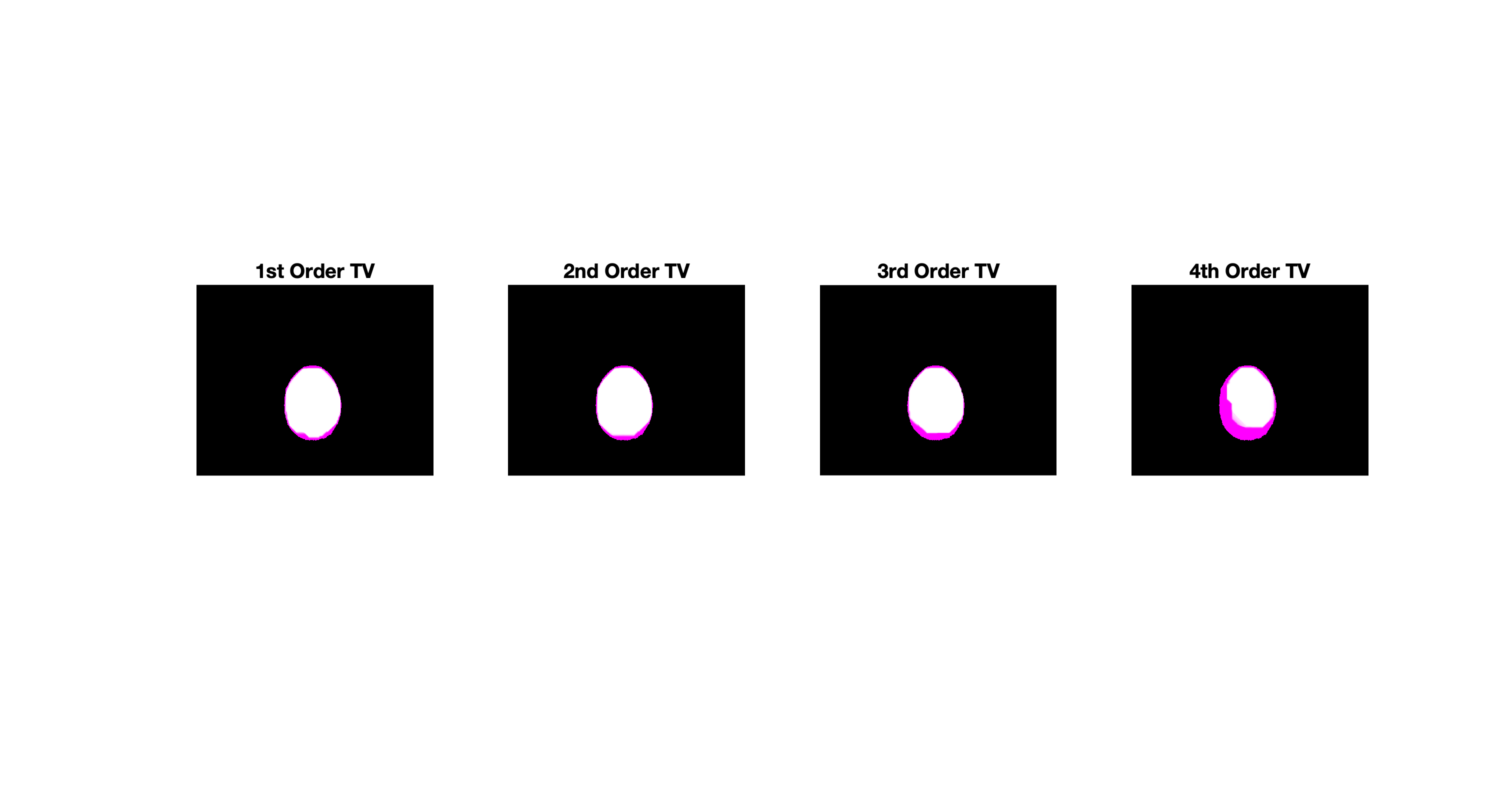
Other fixed values:

= 10

Taylor Series Order = 10

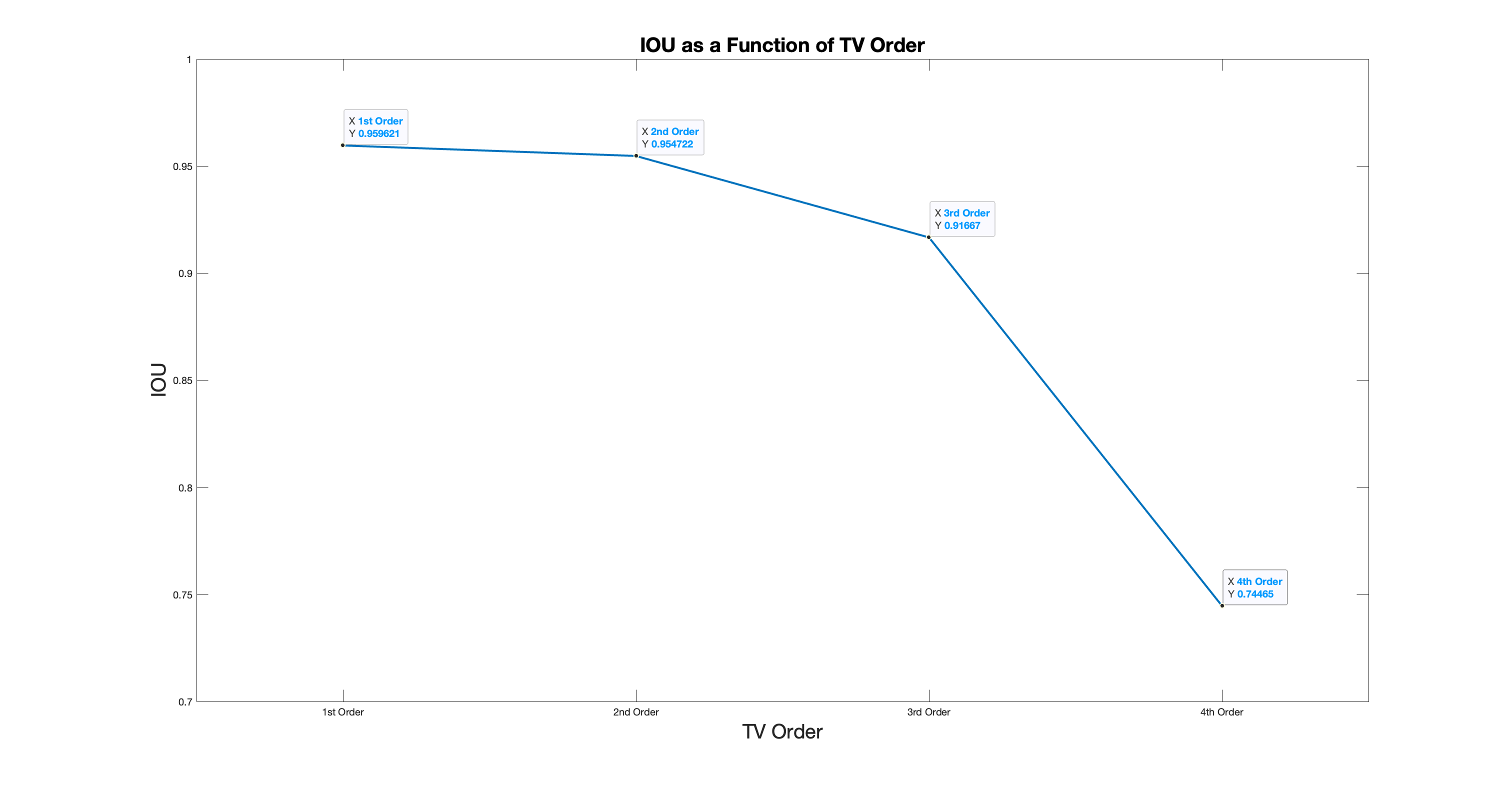
Scale Levels = [16, 8, 4, 2, 1]

The following is the IOU between the algorithm’s mask and ground truth mask in the last image for different TV orders:

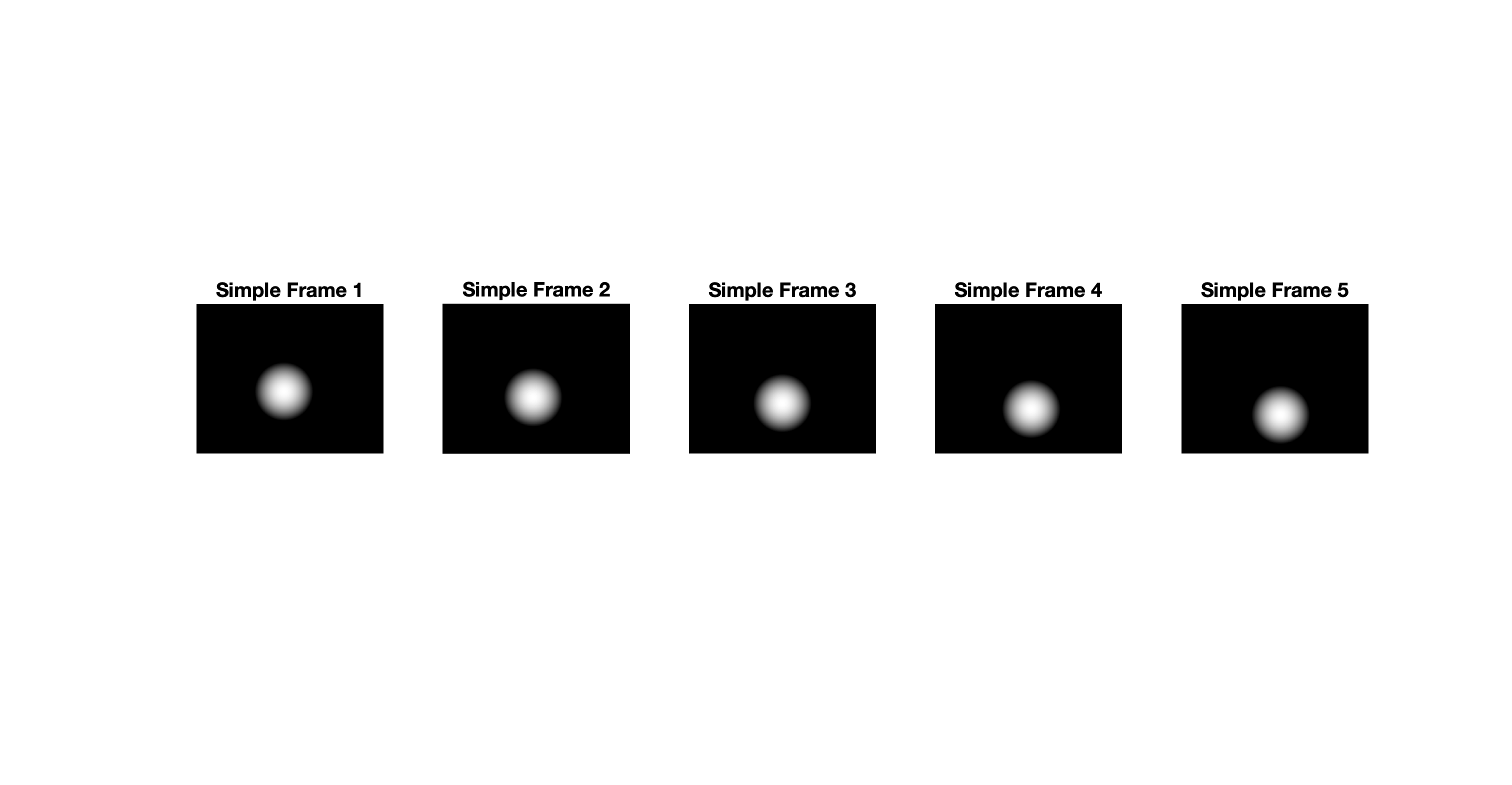


The higher the order the more robust the algorithm is to misalignment of images and its ability to preserve discontinuities, while inducing smoothness. However, as the order gets too high it appears it is too sensitive to discontinuities and produces wrong masks.

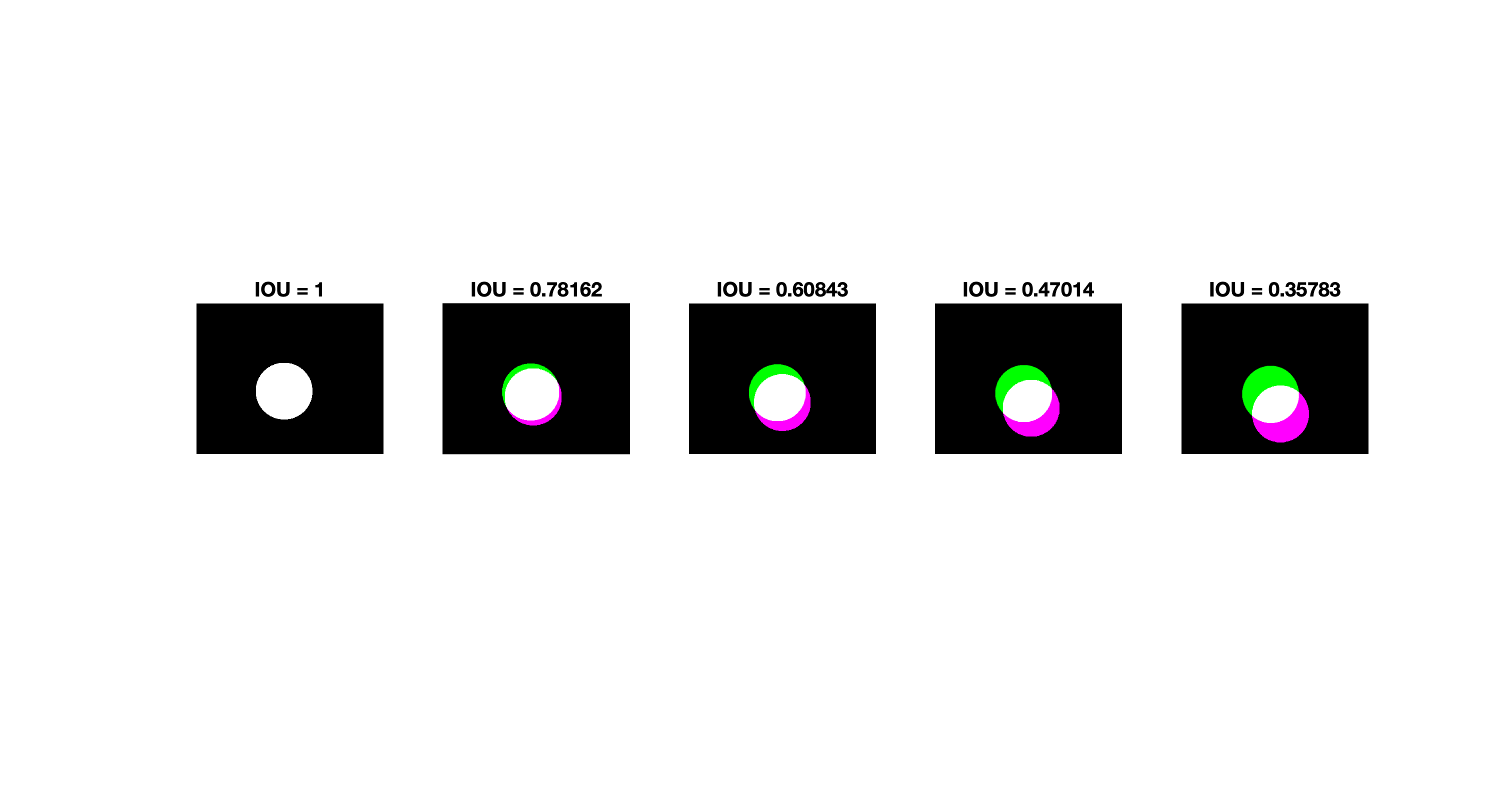
The following is a graph of IOU value as a function of TV order:



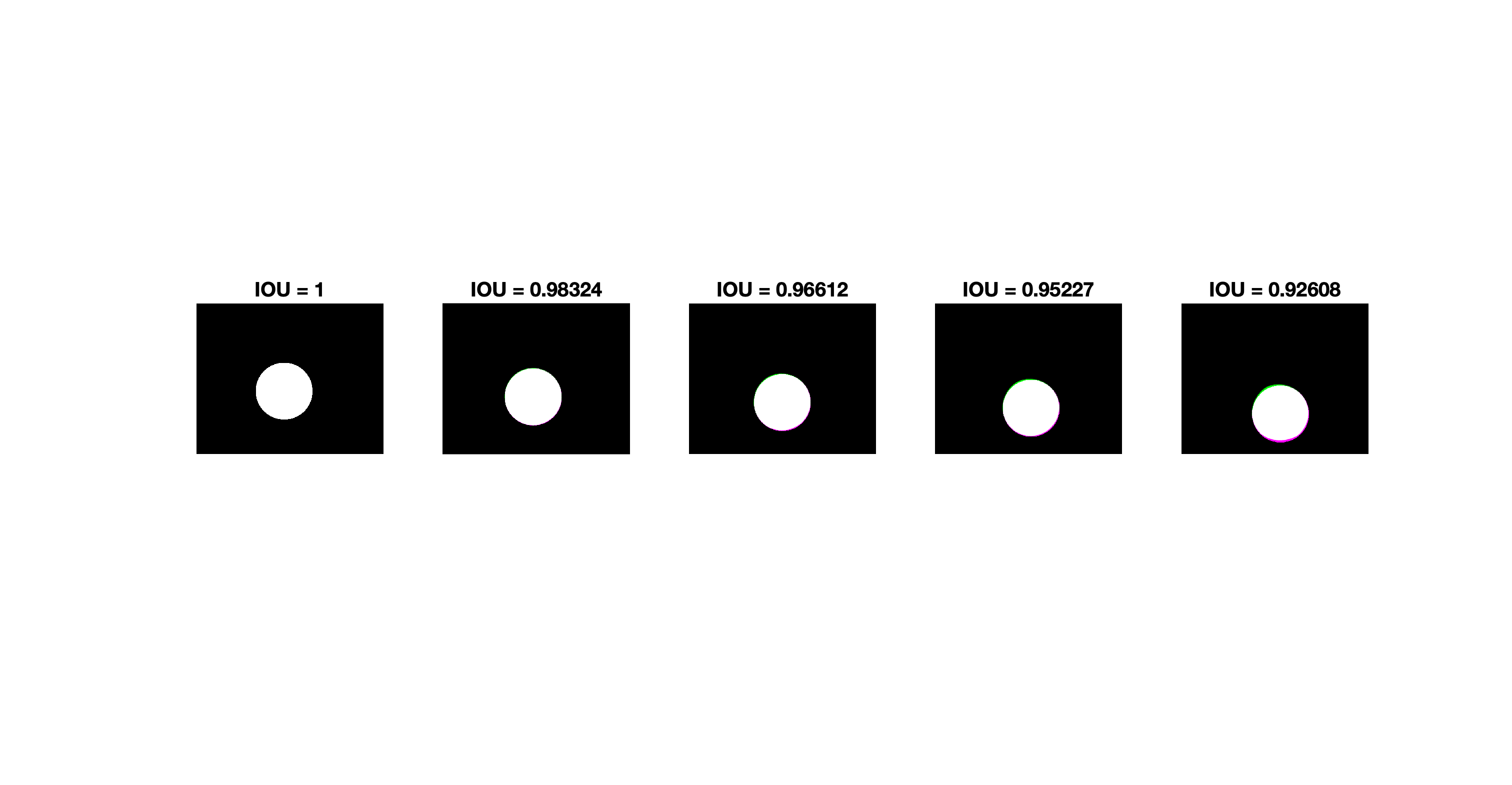
It also seems as if the 1st order TV – Which is regular TV term as learned in class – produced a good mask. However, when a simpler video is used, one where there is only translation between the frames with no change to form:



We can see that 1st order TV is left lacking:



While 2nd order TV handles the translation better:



The data term alone is not enough to reach correct translation through gradient descent. 1st order TV is also not enough as a regularization term.   
As mentioned by the paper, TV suffers when the frames misalign by an affine translation. This is such a case where using higher order TV, which handles the translation successfully, showcases this point.

# Summary

The algorithm, based on the Arbitrary TV method from the paper, successfully functions as an object tracker.

It manages to handle an object’s form changing over time and fix deformities in its mask.

However, there is no correct-works-for-every-scenario set of parameters. Videos with different characteristics need different TV orders, values, etc.

The algorithm is also very slow. Several minutes until convergence between frames. Which makes it non-feasible as a real-time tracker.

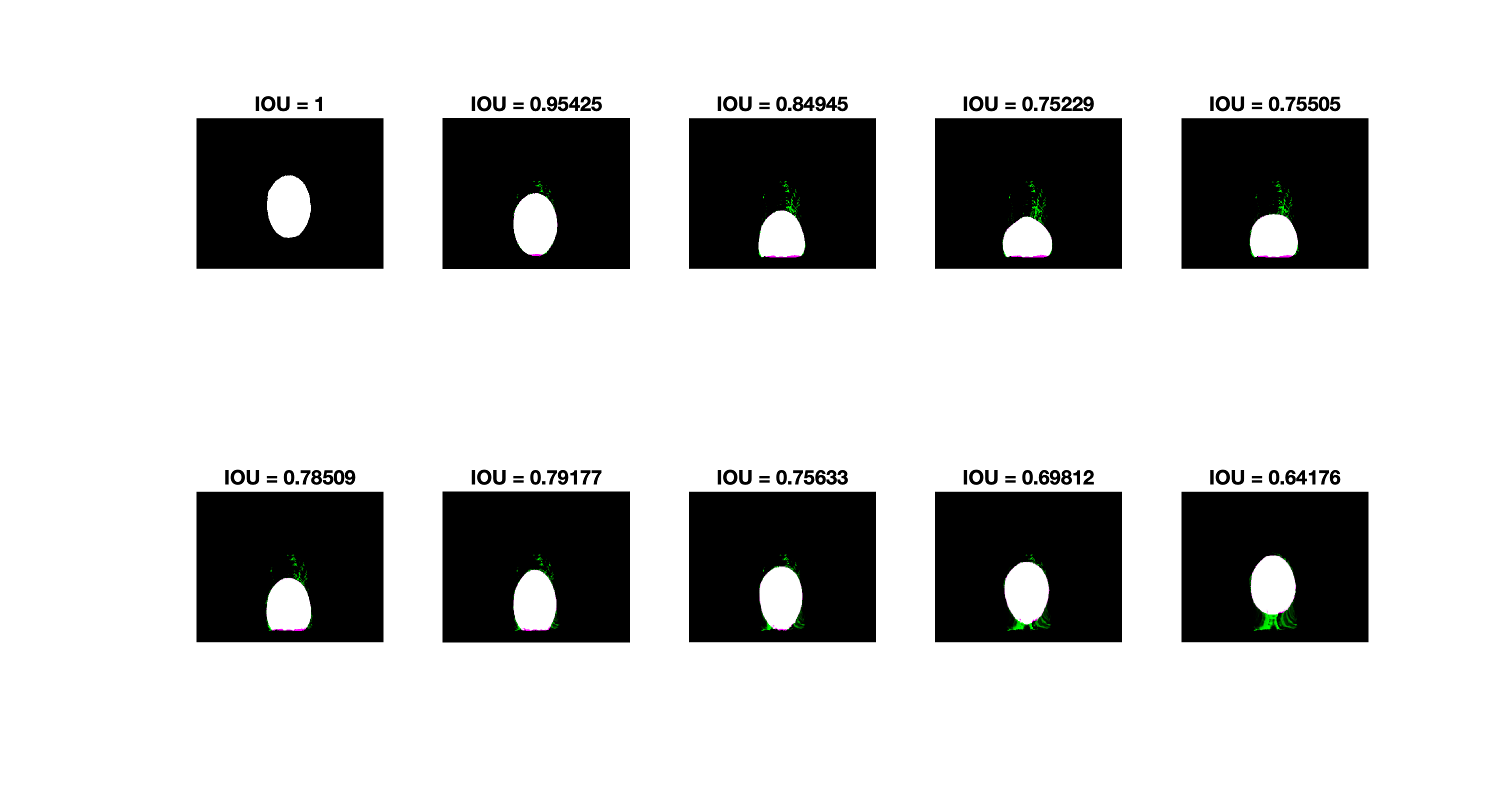
The field of tracking algorithms is already very competitive with the superior technology of neural networks which do work in real-time.

Therefore, in our opinion, the strong competition together with its faults makes our algorithm an interesting prospective, but not one that can reach SOTA status and tour conferences.

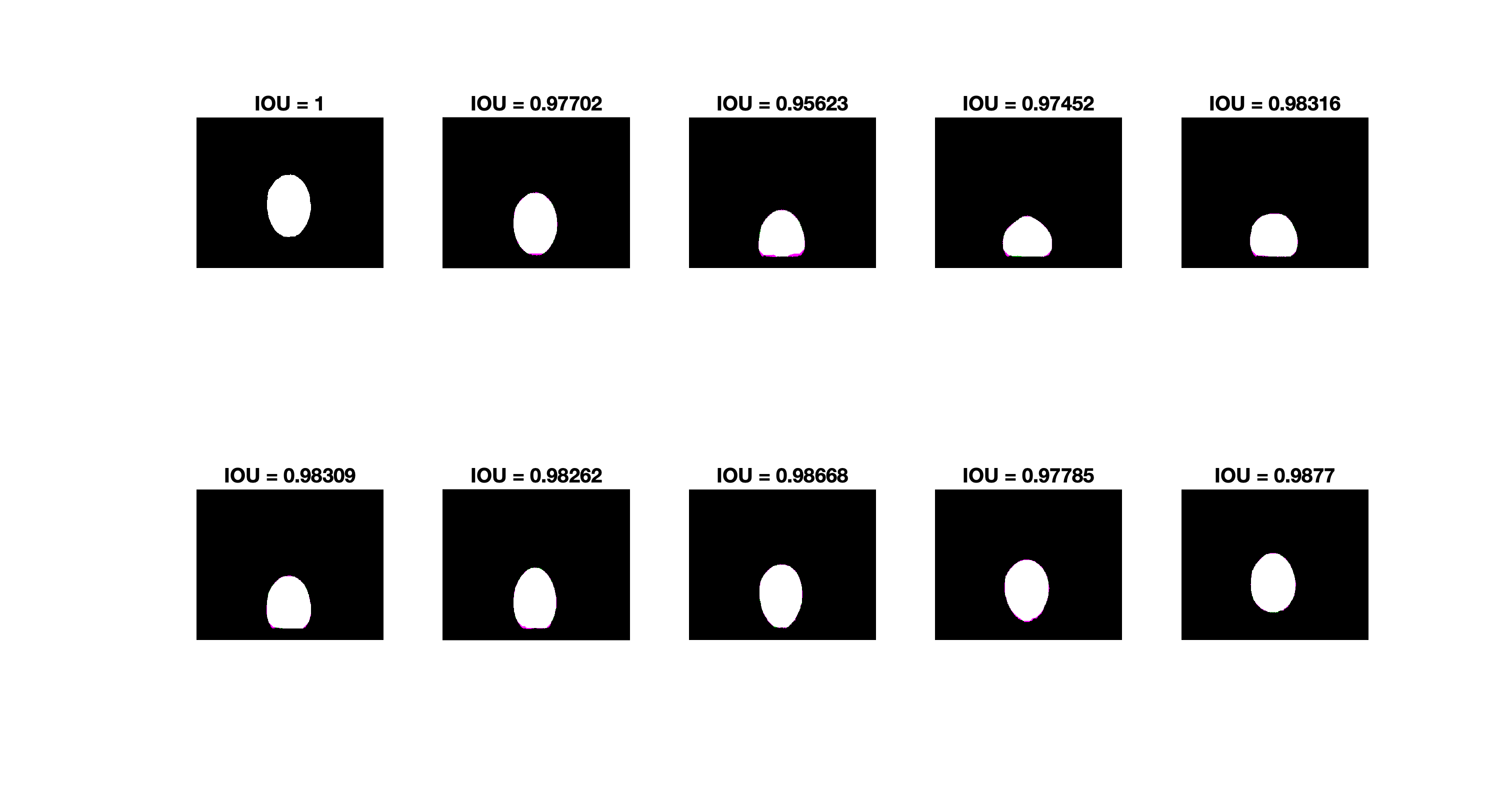
**Effect ofMorphological Operations**

In the algorithm, after the mask is transformed via the displacement field, we apply a morphological operation of **Open** on it (An **Erosion** followed by **Dilation**). Its effect is removing small protrusions and deformities in the mask. We also note that the morphological operation of **Close** might have been useful if small holes in the mask were to appear, which they haven’t.

The following set of images show the results of the algorithm using regularization of 1st order TV with no morphological operations:



As opposed to 1st order with Open operation:



TV as a regularization term benefits the algorithm by making it robust to change in form. However, that same sensitivity can be detrimental, such as in the Tennis ball example – where dust and hairs mix and flow together ending with hairline deformities in the mask. These the morphological operation can easily dispatch.