

1 Algorithms for Segmentation

1.1 Algorithm 1: OpenCV

Image Segmentation with OpenCV

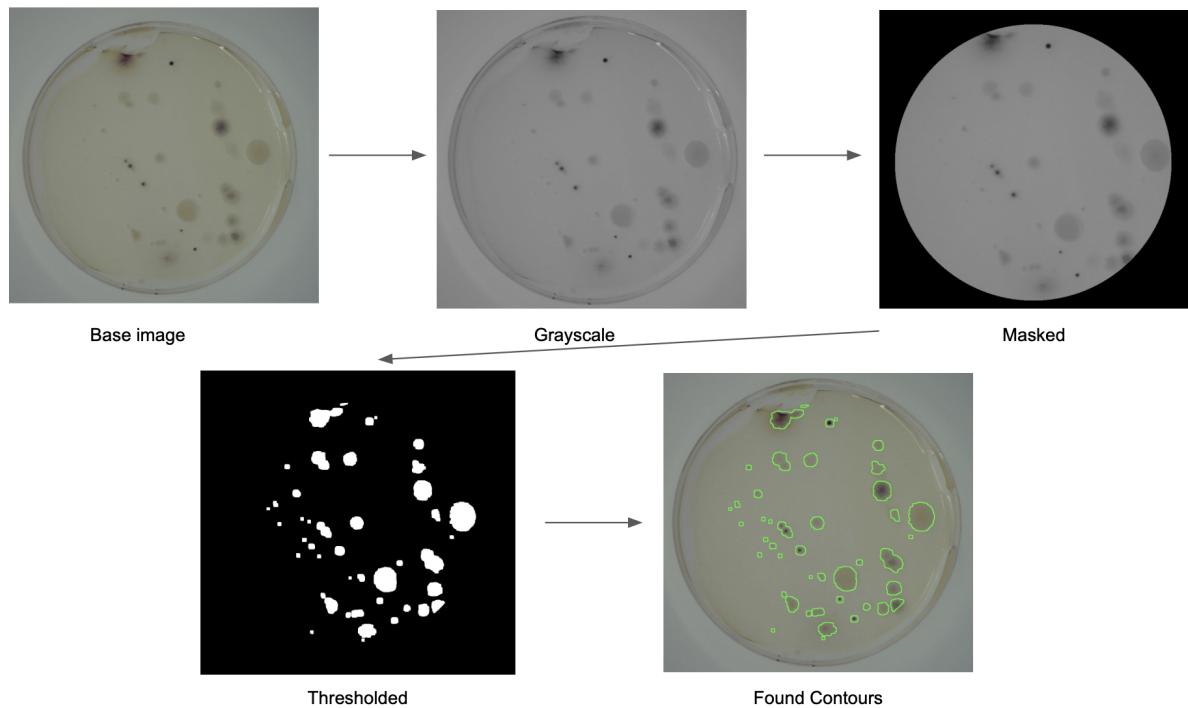


Figure 1: This flow chart represents the algorithm for finding the colonies in OpenCV. The first operation is converting the image into grayscale. This makes thresholding possible and finding the contours easier. Next we find the largest circle (the dish) and use it as a mask. This prevents pixels outside of the dish being detected as colonies. Next we use adaptive thresholding (inverted) to find the colonies in a binary image. From here we use OpenCV to detect the contours within this binary image. These contours are overlaid upon the original image in our final result.

1.2 Algorithm 2: SAM Model

Image Segmentation with SAM Model

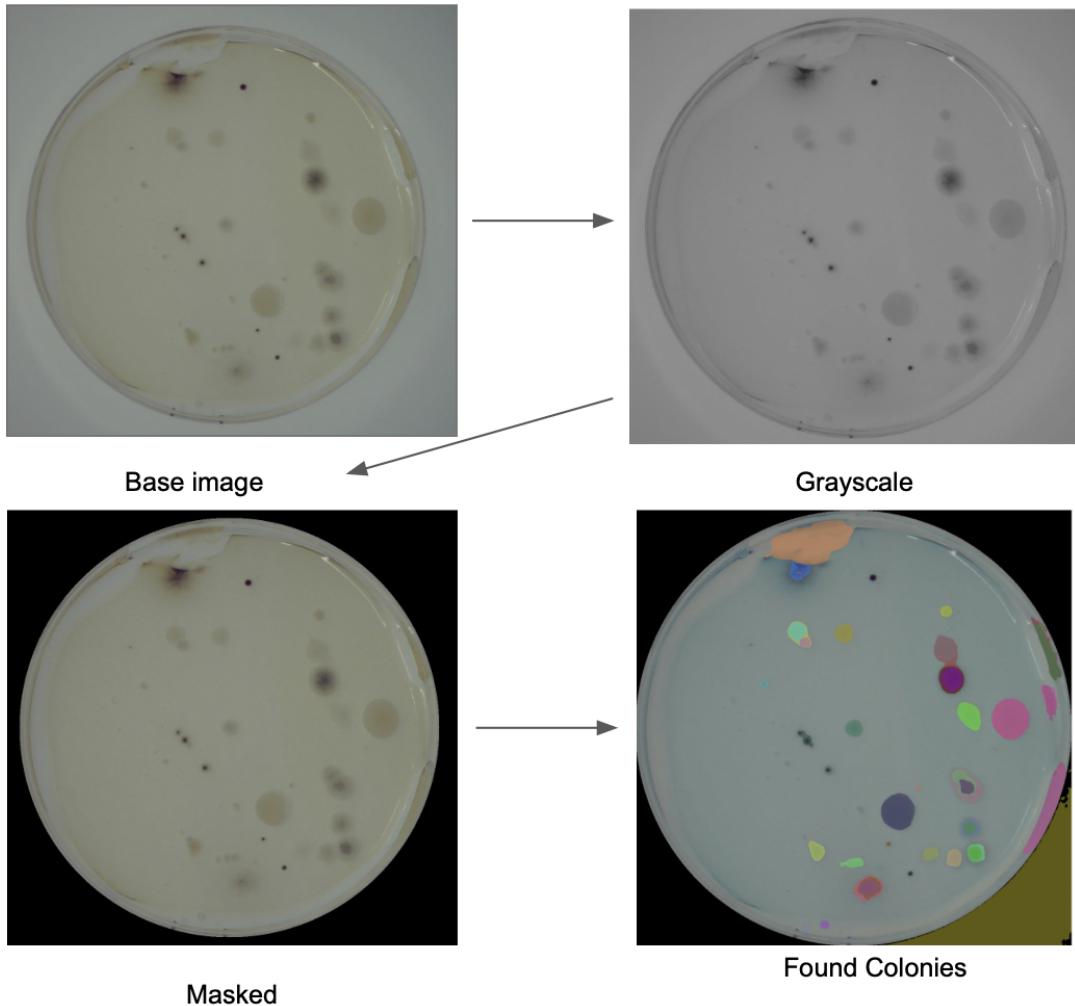


Figure 2: This flow chart represents the algorithm for finding colonies with the SAM (Segment anything Model by Meta) model. The first operation is converting the image into grayscale. This makes thresholding possible and finding the contours easier. Next we find the largest circle (the dish) and use it as a mask. This tries to prevent the dish from being detected as a colony, but does not work very well. After masking the image, we plug it into the SAM model and observe the outputs.

1.3 Algorithm Design Conclusions

The OpenCV model has many advantages over the SAM model. The OpenCV model can be tweaked easily while the SAM model is pre-trained, making it hard to change. Additionally, the OpenCV model is much quicker compared to the SAM model, taking only a couple seconds compared to up to a minute when the image is fed to the SAM model. However, as seen later, the SAM model has certain advantages over the OpenCV model, especially robustness against strange data and overlapping colonies.

1.4 Errors

The SAM model was difficult to constrain to a certain area of the image, so even though I masked the image, I still got colonies counted in the corner. Both models also detected peeling of the solution on the plate as colonies, when they aren't actually colonies.

2 Quantity

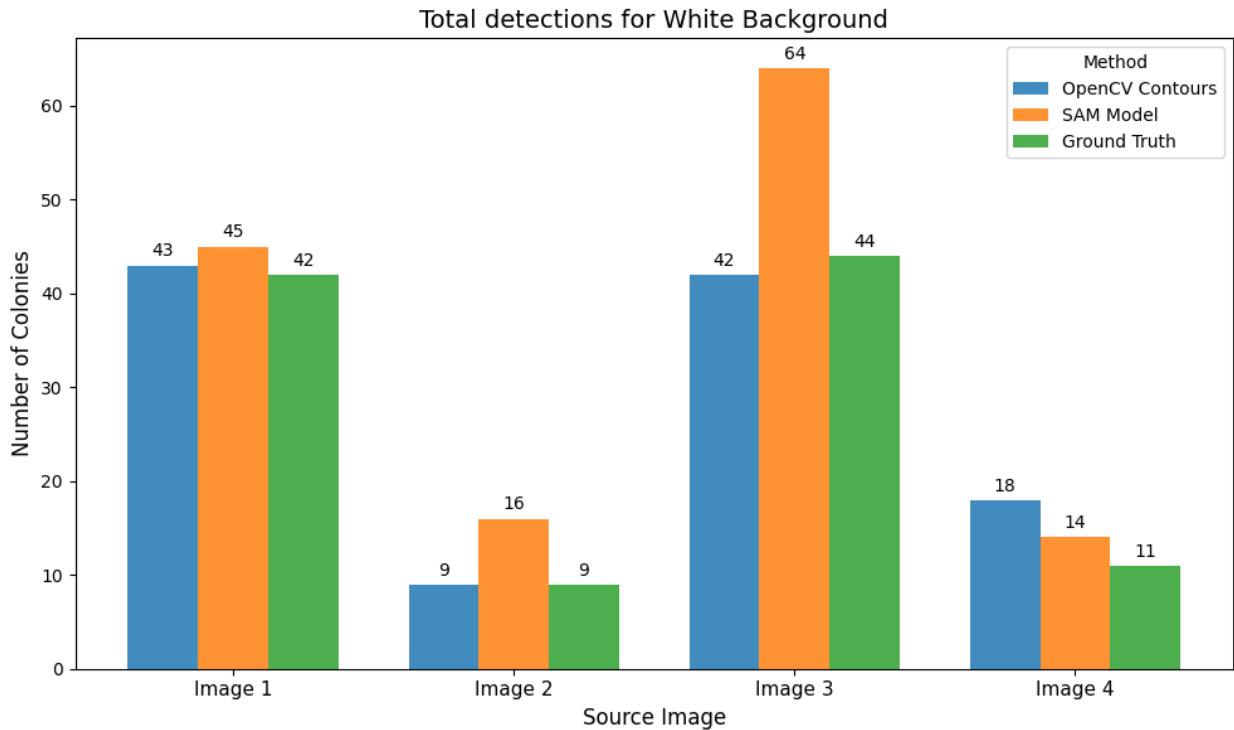


Figure 3: This bar graph shows the difference of quantity of colonies detected by the different algorithms for each image. The tests are performed on the backlit background. The OpenCV method is shown in blue, the SAM model is shown in orange, and the ground truth is shown in green.

2.1 Quantity Conclusions

This figure allows us to see that the SAM model often overestimates the number of colonies, while the OpenCV model is fairly accurate. The only major difference for the OpenCV model is on image 4.

2.2 Quantity Errors

As mentioned earlier, the SAM model detects the edges of the images often, which could be the cause of over-counting. Additionally, the ground truth did not take into account possible colonies that are inside each other, which happened in a couple of the images.

3 Quality

We will gauge quality by the distribution of the areas of the colonies.

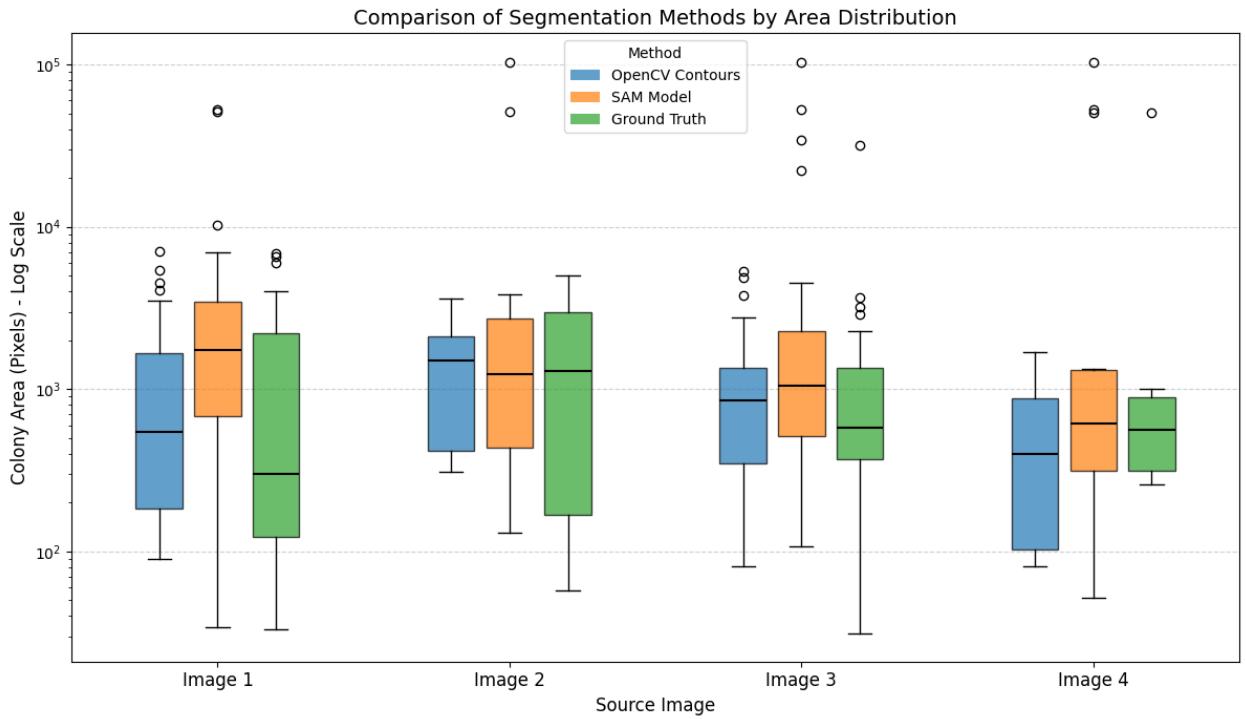


Figure 4: This figure shows the difference in distributions of colony areas between the different segmentation models and the ground truth. The OpenCV model is shown in blue, the SAM model is shown in orange, and the ground truth is shown in green.

3.1 Quality Conclusions

This figure allows us to see that the SAM model often overestimates the sizes of the colonies, with many high area outliers, while the OpenCV model seems to not detect the largest of colonies, failing to see the extremely large outliers. Both models seem fairly ok with their area distributions, but neither of them are particularly good.

3.2 Quality Errors

As mentioned before, the SAM model incorrectly detects the corners of the image as colonies, which may be the cause of the large outliers in the data. Additionally, the smaller colonies in image 4 might be from colonies inside larger colonies in image 4, which isn't represented in the ground truth (as mentioned before). The OpenCV model also underestimates the size of many colonies. This is probably because the algorithm breaks up many of the larger colonies into smaller ones because the adaptive thresholding does not work well on large patches of colonies.

4 Background Comparison

We compare the backgrounds via the count of detected colonies.

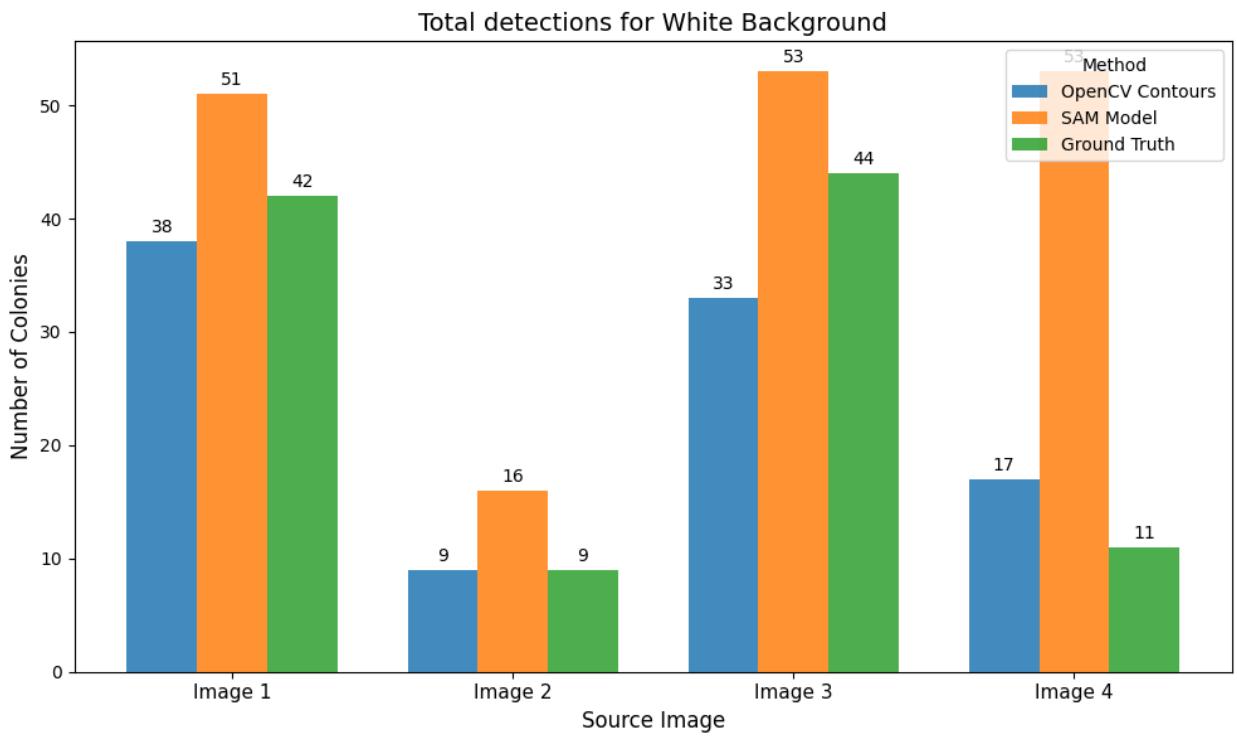


Figure 5: This figure shows the difference in quantity of colonies detected between the different segmentation models and the ground truth. The tests are performed on the white background images. The OpenCV model is shown in blue, the SAM algorithm is shown in orange, and the ground truth is shown in green.

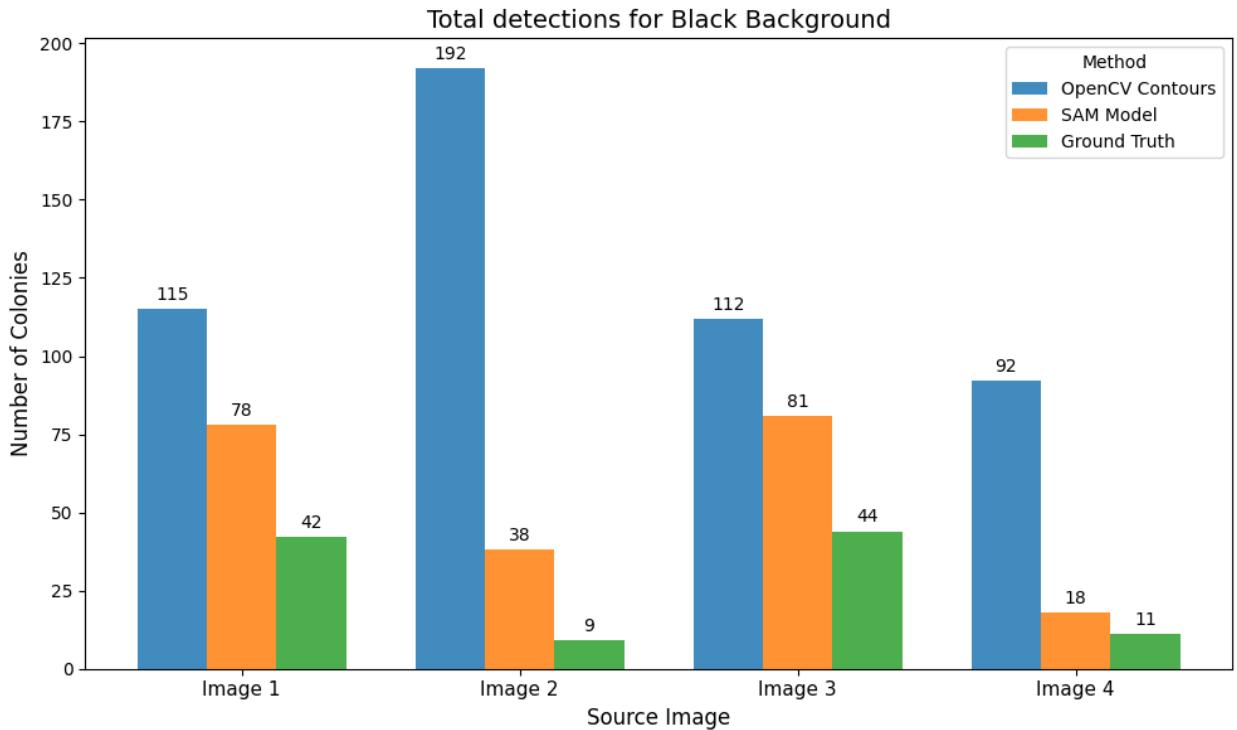


Figure 6: This figure shows the difference in distributions of colony areas between the different segmentation models and the ground truth. The tests are performed on the black background images. The OpenCV model is shown in blue, the SAM algorithm is shown in orange, and the ground truth is shown in green.

4.1 Background Conclusions

The white background significantly messes up the SAM algorithm, especially on image 4. The black background messes up the OpenCV algorithm even more, with image 2 having an incredible overestimate of the number of colonies. These errors make sense, since both models were tuned on the backlit background images, so correct outputs are not expected on the other background types. However, here is also where the strength of the SAM algorithm is shown. This model maintains fairly accurate values in both new environments (barring the one outlier), showing its robustness.

4.2 Background Errors

This section is riddled with errors due to the models being tuned on the backlit images, as stated earlier. Additionally, all the errors from previous sections still apply including the SAM detecting the corners of images and nested colonies.