Experiment 1 - Number of images

- With 5 images the best results are obtained globally
- A minimum of 2-3 images are required to obtain decent results.
- Dreambooth obtains better results than Textual inversion with a smaller number of images.
- A single image can't be used since the results are horrible. Thus, more than one image must be used and that implies that the pictures will contain different subjects of the same class. A situation that is not the main use case of Dreambooth and Textual inversion. Thus, experiment 2 is especially relevant.

Experiment 2 - Different subjects

- Both Dreambooth and textual inversion have an excellent generalization capability for all the cases.
- For different subjects of the same breeds the results are excellent.
- Even for different subjects of different breeds that are similar, both techniques are able to give very good results.
- When the breeds are very different the quality drops dramatically but it is still decent and perfectly recognisable as a dog.

Experiment 3 - Real images percentage

- Accuracy above the baseline for the subject-driven techniques (dreambooth, textual inversion and stable diffusion prompt) when training images percentage below 25%.
- 6/6 tests with training images percentage below 25% show improvements over the baseline.
- For 5% of the training data the biggest improvement is textual-inversion. It goes from 0,6238 to 0,7430. What is a 19,11% improvement.
- For 10% of the training data the biggest improvement is textual-inversion. It goes from 0,7615 to 0,7872. What is a 3,37% improvement.
- Textual inversion is the most promising approach.
- Dreambooth is the most disappointing.
- Classic data augmentation approaches worsen the results when the percentage of real images is low. Nevertheless, when the percentage of training data is bigger than 50% they actually improve the performance of the model.
- RandAugment is the best classic data augmentation technique.
- All classic data augmentation techniques improve the performance when the whole dataset is used.

Experiment 4 - Generated images percentage

- When 100% of real data is used it makes no sense to use subject-driven augmentation. When few images are added the performance is almost the same as the baseline.
- When 100% of real data is used, if a lot of images are added, the performance drops below the baseline with a clear tendency to go even lower.
- When 5% of real data is used it makes a lot of sense to use subject-driven augmentation. When just a few images are added the performance increases dramatically (considering 100% of new images, up to 18,93% improvement with stable diffusion prompt and 11% with textual inversion).
- When 5% of real data is used if a lot of images are added, the performance increases but only up to a point (19,11% with textual inversion and 1000% of new images added)

Experiment 5 - All images generated

- Only using synthetic images decent results are obtained (up to 0,7201 with Textual inversion and 400 images).
- In general terms, the more images the better. Nevertheless, a slowdown of growth can be seen as we surpass more than 100 images per class.
- The maximum accuracy achieved, although competitive, is lower than what can be obtained with real images. That indicates that the quality of synthetic images is lower than real ones. Meaning that they have less fidelity and detail. Still, there is a lot of useful information that can be extracted out of them that is valid for real world scenarios.
- Textual inversion obtained the best results and thus, its images contain the biggest amount of information.

Experiment 6 - ControlNet

- When 100% of real images are used, adding synthetic images does not seem to improve the baseline significantly.
- When 50% of real images are used, adding synthetic images does not seem to improve the baseline at all.
- When 5% of real images are used, adding synthetic images improves accuracy dramatically (up to 23,47%). Considering 100% of images added, up to 12,57% improvement over the baseline. Considering 2000% of images added, up to 23,47% improvement over the baseline.

Experiment 7 - Combinations

- Combining data augmentation techniques does not improve the performance.
- Only in 3/12 cases the performance was increased. And in 2/12 of those cases it was less than 1% improvement.

Experiment 8 - Segmentation

- When 100% of real images are used the improvement seen over the jaccard score baseline is very small, being this in the range 0,42% - 0,66%.
- When 5% of real images are used the improvement seen is higher, getting up to 0,86%. Still, this improvement is very poor.

Experiment 9 - Food-101

- When 100% of the real images are used, adding synthetic images does not seem to improve the baseline significantly.
- When 5% of real images are used, adding synthetic images improves accuracy up to 10,2%.
- The experiment backs up the findings of the rest of the experiments showing that subject-driven augmentation techniques only make sense when the dataset size is small.