# Image Segmentation – Milestone-2

This week, I have completed the EDA and preprocessing of images data. The dataset is published as part of Visual Object Classes Challenge 2012 (VOC2012). The dataset contains 1464 images under the training dataset and 1449 images under the validation dataset. The images are of different shapes. There are two types of images for the segmentation dataset 1. Input images as jpeg and labels as a png image. The label images define the outlined mask of the segments.

## Any surprises from your domain from these data?

The jpeg images are with 3 channels and png images are with one channel. The png image contains the channel value as the label value as per the object. For example, pixel indices correspond to classes in alphabetical order (1=aeroplane, 2=bicycle, 3=bird, 4=boat, 5=bottle, 6=bus, 7=car , 8=cat, 9=chair, 10=cow, 11=diningtable, 12=dog, 13=horse, 14=motorbike, 15=person, 16=potted plant, 17=sheep, 18=sofa, 19=train, 20=tv/monitor). Index 0 is used for background and 255 is used for void or unlabelled. So, while doing augmentation we need both input and label images in the same shape. So, I had to add one more axis before augmenting the mask/label images and remove that axis after augmentation.

## The dataset is what you thought it was?

The dataset I thought will be enough for training a model; however, the training dataset contains only 1464 images. So, if we use a complex model there is more chance of the model getting overfitted.

## Have you had to adjust your approach or research questions?

No, I have not adjusted my research question. However, I am not sure about how much accuracy I can achieve with the pre-trained UNET model.

## Is your method working?

I have not faced any issue other than the augmentation issue about label images mentioned above. I am planning to use vertical flip and horizontal flip augmentation to avoid the overfitting of a model.

## What challenges are you having?

The main challenge with Image data is the size of the data. Currently, I am using 256 X 256 shapes for images. With the hardware available to me, it will take a lot of time to train and tune hyperparameters for my model.