

The creation and intended use of this dataset is described in the paper: “Data Synthesis Methods for Semantic Segmentation in Agriculture: a Capsicum annum Dataset.” Please find the DOI in the description of this dataset. The structure and contents of the dataset is as follows:

/empirical\_image\_color

50 color images of real greenhouse photographs.

/empirical\_label\_class\_colorscale

50 color labels of the empirical\_image\_color set.  
Labels are color coded, intended for convenient human use.  
For computer processing, use empirical\_label\_class\_grayscale instead.

/empirical\_label\_class\_grayscale

This folder consists of subfolders, each with 50 grayscale class labels that correspond to the empirical\_image\_color set.

/empirical\_label\_class\_X\_grayscale

50 grayscale images for class X, coded as RGB = (X,X,X) and (0,0,0) being background.

/empirical\_label\_class\_X\_grayscale\_binary

50 grayscale images for class X, coded as RGB = (1,1,1) and (0,0,0) being background.

/empirical\_label\_class\_all\_grayscale

50 grayscale images with each class C included, coded per class as RGB = (C,C,C) and (0,0,0) being background.

/synthetic\_image\_color

10500 synthetic color images.

/synthetic\_label\_class\_colorscale

10500 color labels of the synthetic\_image\_color set.  
Labels are color coded, intended for convenient human use.  
For computer processing, use synthetic\_label\_class\_grayscale instead.

/synthetic\_label\_class\_grayscale

This folder consists of subfolders, each with 10500 grayscale

class labels that correspond to the synthetic\_image\_color set.

/synthetic\_label\_class\_X\_grayscale

10500 grayscale images for class X, coded as RGB = (X,X,X)  
and (0,0,0) being background.

/synthetic\_label\_class\_X\_grayscale\_binary

10500 grayscale images for class X, coded as RGB = (1,1,1)  
and (0,0,0) being background.

/synthetic\_label\_class\_all\_grayscale

10500 grayscale images with each class C included,  
coded per class as RGB = (C,C,C)  
and (0,0,0) being background.

/synthetic\_label\_depth\_colorscale

10500 depth labels of the synthetic\_image\_color set.  
Images are color coded intended for convenient human viewing.  
For computer analysis, use synthetic\_label\_depth\_grayscale instead.

/synthetic\_label\_depth\_grayscale

10500 depth labels of the synthetic\_image\_color set.  
Images are grayscale coded intended for computer processing.  
Please be referred to the paper mentioned previously for intended use.

/source\_scenes

6 source scenes, each with 7 unique plants. For each plant  
position, 250 frames can be rendered to color, label and depth images.  
Hence, for each source file, 1750 images can be produced, totalling to  
the corresponding 10500 images in the dataset.